USE OF NAVAL FORCE IN CRISES:

A THEORY OF STRATIFIED CRISIS INTERACTION

VOLUME I

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF POLITICAL SCIENCE

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
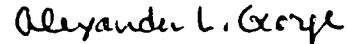
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

by

Joseph Frederick Bouchard

December 1988
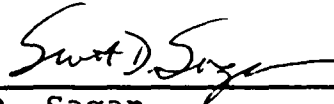
I certify that I have read this thesis and
that in my opinion it is fully adequate, in
scope and quality, as a dissertation for the
degree of Doctor of Philosophy.

Alexander L. George
Alexander L. George
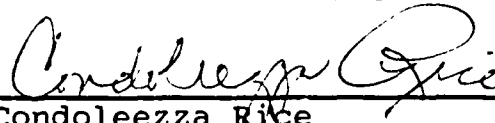(Principal Advisor)

I certify that I have read this thesis and
that in my opinion it is fully adequate, in
scope and quality, as a dissertation for the
degree of Doctor of Philosophy.
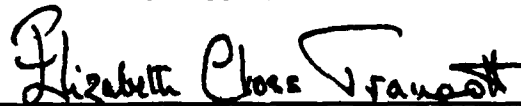
Scott D. Sagan
Scott D. Sagan

I certify that I have read this thesis and
that in my opinion it is fully adequate, in
scope and quality, as a dissertation for the
degree of Doctor of Philosophy.

Condoleezza Rice
Condoleezza Rice

Approved for the University Committee
on Graduate Studies:

Elizabeth Close Vrancott
Dean of Graduate Studies

iii

USE OF NAVAL FORCE IN CRISES:
A THEORY OF STRATIFIED CRISIS INTERACTION

Joseph Frederick Bouchard, Ph.D.
Stanford University, 1988

Previous studies of international crises have implicitly viewed all of the political and military interactions between the two sides as a single interaction sequence. This fails to capture the complexity of crisis interaction and crisis stability. The theory of stratified interaction developed in this dissertation states that crisis interaction occurs at three levels: political, strategic, and tactical. Interactions at each level evolve separately and can independently influence whether or not a crisis escalates to war.

The objective is to develop a differentiated theory of crisis interaction cast in the form of contingent generalizations that offer discriminating explanations for the occurrence of crisis stability problems. The method of structured, focused comparison is used to conduct empirical research on two sets of historical cases. The first set consists of four case studies of United States naval operations in crises: the 1958 Taiwan Straits Crisis, the 1962 Cuban Missile Crisis, the 1967 Arab-Israeli War, and

the 1973 Arab-Israeli War. The second set consists of four
case studies of peacetime attacks on United States Navy
ships: the 1964 Tonkin Gulf Incidents, the 1967 Israeli
attack on the USS Liberty, the 1968 North Korean seizure of
the USS Pueblo, and the 1987 Iraqi attack on the USS Stark.

Structured comparison of these cases reveals that the
stratified interaction model provides an accurate descrip-
tion of international crises. Tactical-level interactions
normally are not under the direct direct control of national
leaders, and under certain conditions can become decoupled
from the political-military objectives and strategy of
national leaders. Five patterns of tactical-level interac-
tions are identified: parallel stratified interactions,
momentary decoupling, decoupling followed by disengagement,
decoupling followed by tactical-level escalation, and
decoupling causing escalation at the strategic or political
levels. The factors that can cause decoupling of tactical-
level interactions and the factors that determine whether or
not decoupled interactions escalate uncontrollably to war
are identified. Additionally, three political-military
tensions that can arise in crises are identified: tension
between the needs of diplomatic bargaining and the needs of
military operations, tension between the need for top-level
control of military operations and the need for tactical-
level flexibility and initiative, and tension between crisis
military operations and readiness for wartime missions.

PREFACE

The opinions expressed in this dissertation are the author's alone and are not to be construed as representing views or policies of the Department of the Navy. All source material cited in this dissertation is unclassified. Portions of this dissertation were submitted for security review in accordance with applicable Department of Defense and Department of the Navy instructions. That review did not result in any changes in the substance of the dissertation, and did not in any way restrict the academic freedom of the author.

My foremost acknowledgement is to the United States Navy, which funded my graduate studies at Stanford University through the Junior Line Officer Advanced Education Program (Burke Scholar Program) and provided me time off from my Navy duties to pursue the degree. The scholarship program was managed by the Civilian Institutions Program at the Naval Postgraduate School, Monterey, California. The staff of the Civilian Institutions Program provided superb support for me while I was at Stanford.

I am particularly appreciative of the Stanford University Department of Political Science graduate program

assistants, Jule Kringle and Dorothy Blake. Both were marvelously patient and persevering in dealing with the unique problems of having an active duty naval officer as a graduate student. The Center for International Security and Arms Control at Stanford University, directed by Professor Jown W. Lewis, accepted me as a fellow for two years, providing a stimulating academic environment.

The Naval Historical Center, Washington, D.C., provided invaluable assistance for my research. I am particularly indebted to Bernard Cavalcanti, Westly Price, and the staff of the Operational Archives Branch. The Ship's History Branch and Naval Aviation History Branch and the Department of the Navy Library were also of assistance.

I would like to thank several other individuals who assisted my research: Commander Figueras in the Office of the Chief of Naval Operations (OP-602D), the staff of the Navy Command Center, Vice Admiral Jonathan T. Howe at the Joint Chiefs of Staff, The Honorable William A. Cockell, Jr., at the National Security Council, and Paul Stillwell and Susan Sweeney at the U.S. Naval Institute's Oral History Program. Among the naval officers who read and critiqued portions of my dissertation, Captain Robert A. Doll was especially helpful. I thank Lieutenant Commander Yarosh and the Administrative Office at Naval Air Station Moffett Field for handling the extra administrative burdens created by my graduate program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Figure

CHAPTER I

INTRODUCTION

Studies of international crises have repeatedly concluded that the success of crisis management efforts is critically dependent upon top-level political authorities maintaining close control of the actions of their military forces. This essential crisis management requirement has been identified as a potentially serious problem area.

Several concerns have been raised: Preplanned military operations and contingency plans may not be appropriate for the unique circumstances of a particular crisis, and may not support the political-diplomatic strategy adopted by national leaders to resolve a crisis. Delegated command of military operations could allow unintended military incidents to occur, which the adversary could misperceive as a deliberate escalation of the crisis or signal of hostile intent. Military alerts ordered to deter the adversary and increase the readiness of the armed forces could set in motion a chain of events exceeding the control of national leaders. Such problems are sources of concern because they could cause national leaders to lose

1

control of events in a crisis, starting an escalatory spiral leading to war.[1]

On the other hand, some scholars believe that while inadvertent military actions can contribute to crisis management problems and the occurrence of inadvertent war, attention should be focused on the political and psychological pressures on top-level decisionmakers. Thomas C. Schelling, in a passage particularly relevant to this study, has expressed such a view:

> This is why there is a genuine risk of major war not from "accidents" in the military machine but through a diplomatic process of commitment that is itself unpredictable. The unpredictability is not due solely to what a destroyer commander might do at midnight when he comes across a Soviet (or American) freighter at sea, but to the psychological process by which particular things become identified with courage or appeasement or how particular things get included in or left out of a diplomatic package.[2]

Thus, there is disagreement among students of crisis and war over the effects of inadvertent military incidents on crisis

---

[1] For examples of such concerns, see Alexander L. George, David K. Hall, and William E. Simons, The Limits of Coercive Diplomacy (New York: Little, Brown and Co., 1971), p. 14; John Steinbruner, "An Assessment of Nuclear Crises," in Franklin Griffiths and John C. Polanyi, eds., The Dangers of Nuclear War (Toronto: University of Toronto Press, 1980), p. 40; Phil Williams, Crisis Management (New York: John Wiley, 1976), p. 202; Richard N. Lebow, Between Peace and War (Baltimore: Johns Hopkins University Press, 1981), p.287; Alexander L. George, "Crisis Management: The Interaction of Political and Military Considerations," Survival 26 (September/October 1984), pp. 227-228; Scott D. Sagan, "Nuclear Alerts and Crisis Management," International Security 9 (Spring 1985): 99-139.

[2] Thomas C. Schelling, Arms and Influence (New Haven, CT: Yale University Press, 1966), p. 93.

stability. Inadvertent military incidents are viewed as dangerous in and of themselves because they can directly trigger escalation, or, alternatively, are viewed as dangerous because of the manner in which they can influence the perceptions held by national leaders. This issue is of practical relevance in crisis decisionmaking, as well as being of theoretical interest in the study of international crises.

The focus of this study is on the problems that can arise when using military force as a political instrument in crises. In an international crisis, military forces commonly perform two missions: political signalling in support of crisis bargaining, and preparing for localized fighting and war should crisis management efforts fail. Inadvertent escalation--any increase in the level or scope of violence in a crisis that was not directly ordered by national leaders or anticipated by them as being the likely result of their orders--is a significant danger in these circumstances. A distinction can be drawn between the general political requirements of crisis management, such as limiting political objectives and military means, and the operational requirements of crisis management, such as maintaining control of military operations.[3] The focus of this study is on the operational requirements of crisis

---

[3]This distinction pointed out to the author by Alexander L. George.

management.[4] The use of United States naval forces in four crises that occurred since the end of World War II will be examined to develop contingent generalizations on crisis military interaction.

In the introduction to his study of international crises, Richard N. Lebow discusses the distinction long made between the underlying causes of war, the long-term sources of hostility and tension, and the immediate causes of war, the particular events, such as a crisis, sparking a war. Lebow argues that, while students of international relations since Thucydides have focused on underlying causes, immediate causes are at least as important as underlying causes, in that immediate causes can determine whether or not war erupts from the underlying hostility and tension.[5] This study starts from the premise that Lebow is correct, that immediate causes are important for understanding how and why wars occur.

The causes of war can be viewed as falling on a time-span spectrum, with long-term underlying causes working their effects over years, decades or even centuries toward the left end, and immediate causes occurring over days or weeks toward the right end. The underlying causes toward

---

[4]Scott Sagan refers to this as the operational dimension of crisis management. See Sagan, "Managing Strategic Nuclear Alert Operations," (Santa Monica, CA: Rand, forthcoming 1988).

[5]Lebow, pp. 1-4.

the left end of the spectrum include the structure of the international system, history, culture, economic development and resources, ideology, geography, and military technology. System structure has a strong influence on how "war-prone" international politics are at a given time. Historical, cultural, economic, and ideological variables help to shape the political framework within which rivalries arise between particular nations and contribute to the intensity of the hostility and tensions between them. Geographic factors, and the state of military technology shape the strategic relationships between nations and contribute to the level of tensions between them.

This study will be addressing causes of war at the far right end of that spectrum--events occurring over hours, or even just minutes at the speed of modern warfare. There is no intent to slight the importance of underlying causes or longer-term immediate causes, which arrange the political and strategic circumstances for war to occur. Rather, the intent is to supplement those causes with greater understanding of how military interactions in a crisis could inadvertently trigger war.

One of the fundamental problems in international relations is to identify the necessary and sufficient conditions for war to occur. This study makes two assumptions on the necessary and sufficient conditions for war. The first is that an international environment marked

by confrontation over national interests, hostility, and tension--all arising from the underlying causes of war--are a necessary condition for war to arise from a crisis. The implication of this assumption is that inadvertent military incidents will not spark escalation leading to war in the absence of confrontation, hostility and tensions. This study thus focuses on inadvertent escalation arising under conditions of acute international crises, when the necessary condition for war are present.

The second assumption is that the underlying causes of war are not sufficient conditions for war. War can be avoided even under conditions of confrontation, hostility, and tension so long as national leaders on each side are willing to continue bargaining with the other side, are willing to sacrifice certain interests in order to protect or advance others, and perceive that the other side intends to continue bargaining rather than resort to war. This suggests that a number of factors can provide conditions sufficient for war once the necessary conditions are present. Examples include a belief that vital national interests cannot be protected through bargaining, an unwillingness to concede some interests to protect others (perhaps because the price would be too high or domestic political repercussions too severe), a misperception that the other side will not bargain seriously or intends to resort to war at an opportune moment, and loss of control

over military operations. These factors can give rise to either deliberate decisions to go to war or to inadvertent war. The immediate causes of war can thus provide sufficient conditions for war if the necessary conditions are present. The focus of this study is on a specific subset of the immediate causes of war: those arising from interaction of the military forces of the two sides and resulting in inadvertent escalation to war.

The remainder of this introduction will present a brief critique of the literature on crises and crisis management, an overview of the theory being proposed, the research design, the historical cases and case selection criteria, and a summary of the organization of the study.

Critique of Crisis Theories

The existing literature on crises and crisis management has three serious weaknesses. First, the various political and military interactions that occur between the two sides in a crisis are assessed in the context of an implicit single interaction sequence model of crises. Second, the frequently observed phenomenon of United States leaders exercising close control over military operations in crises, combined with a lack of familiarity with military command and control procedures, has produced an erroneous view of the manner in which military forces are controlled in crises. Third, and derived from the two previous

weaknesses, the concept of crisis stability is poorly developed and there is a poor understanding of the escalation processes that could cause a crisis to escalate to war. These three weaknesses in the crisis management literature are discussed in Chapter II. They are summarized here to provide an overview of the study.

The first weakness is that previous studies of international crises have implicitly viewed the various political and military interactions that occur between the two sides as a single interaction sequence. The flow of events in a crisis is viewed as a single sequence of actions and reactions. A consequence of this perspective is the implicit assumption that all the actions taken by a nation during a crisis either are ordered by national leaders in pursuit of their policy objectives, or should not have occurred and therefore represent a loss of control over events. Under the single interaction sequence model of crisis interaction, a policy objective desirable for avoiding war--control of crisis military operations by top-level political authorities--is treated as the norm against which actual crisis management efforts are compared. The occurrence of military interactions not directly controlled by national leaders is then viewed as a potentially dangerous breakdown of crisis management.

The single interaction sequence model does not accurately describe international crises. What actually

occurs is multiple interaction sequences that only partially influence each other. Multiple interaction sequences, evolving simultaneously but semi-independently, arise when national leaders do not make all operational decisions themselves, but must delegate significant decisionmaking authority to subordinates. This is the basis for the stratified interaction model of international crises, described in detail in Chapter III.

The second weakness in the crisis management literature is that it is based on an erroneous view of the manner in which military forces are controlled in crises. This apparently resulted from the frequently observed phenomenon of United States leaders exercising close control over military operations in crises, combined with a lack of familiarity with military command and control procedures. The crisis management literature typically describes the control of crisis military operations as being highly centralized, with top-level civilian authorities exercising direct control--in contrast to routine peacetime operations, which are described as highly decentralized and having little involvement of civilian political authorities. This description fails to grasp the true complexity of military command and control, leading to inaccurate assessments of the crisis management problems arising from the employment of military forces in crises and how those problems can affect crisis stability.

Even in crises, military commanders are delegated significant authority to make operational decisions on the employment of their forces--including decisions on the use of force. Under certain circumstances military commanders can use conventional weapons without seeking permission from higher authorities. The scope of their authority is spelled out in a variety of documents, which collectively will be referred to as mechanisms of indirect control. There are even provisions for commanders to act contrary to their written instructions when circumstances dictate.

Although some scholars have recognized that these features exist in the United States military command and control system, the actual complexity of that system has not been reflected in the literature on crisis management. The literature is founded on a simple distinction between policy-making and policy implementation, and turns to concepts such as bureaucratic politics and organizational process to explain why actions are taken that were not ordered by national leaders. This fails to recognize that military commanders at all levels in the chain of command have important policy-making roles and are not simply bureaucrats executing policy decisions. Thus, an understanding of the mechanisms through which authority to make operational decisions is delegated to military commanders is essential for accurately assessing the crisis management problems that arise when military forces are employed in crises.

The third weakness in the crisis management literature is that the concept of crisis stability is poorly developed and there is a poor understanding of the escalation processes that could cause a crisis to escalate to war. Crisis stability is viewed as being primarily a function of weapons technology, particularly the degree to which it gives an advantage to the offense. Lacking is an appreciation of the operational factors that affect crisis stability once a decision is made to employ military forces in a crisis. The escalation processes that could cause a crisis to escalate to war are also poorly developed. Although there is growing concern over inadvertent or accidental war, these concepts are not well defined and the scenarios in which they could occur lack plausibility. Crisis military operations can indeed trigger or contribute to an escalatory process leading to war, but the manner in which they do so are subtle and complex--and best understood in the context of stratified interactions.

To summarize, the weaknesses in the crisis management literature are an implicit and misleading single interaction sequence model of the political and military interactions that occur in a crisis, an erroneous view of the manner in which military forces are controlled in crises, and poor development of the concept of crisis stability and the escalation processes that could cause a crisis to inadvertently escalate to war.

Overview of Concepts and Theory

Three central concepts form the foundation for this study: stratified interaction, stratified crisis stability, and the tensions that arise from the interaction of political and military objectives in a crisis. The theory and its corollaries are developed and explained in detail in Chapter III. They are summarized here to provide an overview of the concepts presented in the research design.

The scope of this study is limited to international crises in which two fundamental conditions are present: The first is that both sides in a crisis seek to protect or advance vital national interests and, conversely, have vital interests at stake that they are unwilling to sacrifice for the purpose of avoiding war. Both sides thus take military actions intended to support crisis bargaining and to counter military moves by the other side. The second assumption is that neither side desires war as the outcome of the crisis. National leaders on each side limit their objectives and restrain their military moves to avoid being misperceived by the other side as intending to launch a war. Both sides thus seek to avoid inadvertent escalation of the crisis while deterring escalation by the other side. When both of these conditions are met, the primary danger is of war arising from inadvertent escalation. These conditions and the nature of international crises are discussed in detail in Chapter II.

The theory of stratified interaction states that, given conditions of delegated control, tight horizontal coupling between the military forces of the two sides, and acute crisis, interactions between the two sides will be stratified in three levels: political, strategic and tactical. The first corollary to the theory is that tactical-level interactions can become decoupled from the political-military objectives of national leaders. The term decoupled is used to mean that vertical command and control links to operational military forces at the scene of a crisis are severed or otherwise fail to ensure that tactical-level decisionmaking supports the crisis management strategy of national leaders. Decoupling occurs to the extent that operational decisions on the employment of military forces made at the strategic and tactical levels differ from the operational decisions political level decisionmakers would have made to coordinate those military actions with their political-diplomatic strategy for resolving the crisis. This is an inductive theory arrived at through empirical historical research into crisis interactions.

Crisis stability exists to the extent that neither side has an incentive to strike the first military blow. The crisis security dilemma is that, in a crisis, many of the actions a state takes to increase its security and improve its bargaining position decrease the security of the adversary. The theory of stratified interaction directly

affects this dilemma. The stratified crisis security dilemma is that, in a crisis, the security dilemma is stratified, arising from the interaction processes occurring separately at each of the three levels, and affecting the likelihood of violence separately at each level. This in turn leads to the concept of stratified escalation dynamics: in an acute crisis, in which tactical-level interaction between the two sides has become decoupled from direct control by national leaders, the security dilemma, operating separately at the tactical level, can trigger an escalatory spiral, which under certain circumstances can cause the crisis to escalate uncontrollably to war. Identifying those circumstances is a primary objective of this study.

An important issue is whether these phenomena—stratified interaction, decoupling of tactical-level interactions, and stratified crisis stability—are strictly symmetrical or can also be asymmetrical. That is, must the conditions necessary for these phenomena to occur be present on both sides in a crisis, or can can the phenomena arise when the conditions are present on only one side. This issue will be addressed in the empirical research on the theory, but the focus of the study will be on the United States and the role of U.S. forces in crises. The preliminary assessment is that stratified interaction tends to be symmetrical (both sides in a crisis normally experience the conditions for stratification), but that

decoupling and stratified crisis stability can be either symmetrical or asymmetrical. This has important implications for crisis management: war could arise through a process in which one side has lost effective control of its forces and is experiencing inadvertent escalation, while the other side retains control over its forces and is deliberately escalating the level of violence. This is probably a more likely and dangerous scenario than one in which symmetrical decoupling occurs and both sides experience inadvertent escalation.

Another aspect of crisis stability is the danger of misperception under conditions of stratified interaction. The concept of the misperception dilemma describes the inadvertent results that can occur when military forces are used for signalling in a crisis. When signalling adversaries, the dilemma is between inadvertent signals of hostility and inadvertent signals of acquiescence. When signalling an ally or friend, the misperception dilemma is between inadvertent signals of encouragement and inadvertent signals of retrenchment. Given stratified interactions, then perceptions of the adversary can also be stratified, with different perceptions being held at different levels of interaction. Misperceptions can arise at one level without other levels necessarily being aware of them, providing a mechanism by which stratified interactions can become decoupled.

The interaction of political and military considerations when military force is employed as a political instrument in crises will be a central focus of the study. The interactions generate what will be described as political-military tensions--actual and potential conflicts between political and military considerations which force decisionmakers, either knowingly or tacitly, to make trade-offs among individually important but mutually incompatible objectives. These political-military tensions, which can give rise to difficult policy dilemmas in a crisis, are inherent in the use of force as a political instrument under conditions of stratified interaction.

There are three political-military tensions. The first is tension between political considerations and the needs of diplomatic bargaining, on the one hand, and military considerations and the needs of military operations, on the other. The second is tension between the need for top-level control of military options in a crisis, and the need for tactical flexibility and instantaneous decisionmaking at the scene of the crisis. The third is tension between performance of crisis political missions and readiness to perform wartime combat missions. These three tensions between political and military considerations affect the degree to which stratified interactions become decoupled in a crisis, thus having a significant impact on crisis decisionmaking and crisis stability.

Research Design

There is an inherent element of randomness and unpredictability in the occurrence of war that structural or system-level theories cannot eliminate or define out of existence. Addressing the immediate causes of war gets at that element of randomness and unpredictability, allowing identification of various sets of specific circumstances in which the probability of war is increased--which is both theoretically significant and policy relevant. This study will examine a particular subset of the immediate causes of war, those arising from the use of force as a political instrument in crises.

The type of theory this study seeks to develop is what Alexander L. George describes as a "differentiated" theory, an explanatory theory cast in the form of contingent generalizations and offering discriminating explanations for the occurrence of a phenomenon. Contingent generalizations are regularities that occur only under certain specific conditions. Collectively they offer a differentiated typology of situations in which the phenomenon of interest manifests itself. The objective of a differentiated theory is to identify the variety of different causal patterns that can occur for the phenomenon, and the conditions under which each distinctive causal pattern occurs. The value of a differentiated theory is that it has greater policy

relevance than theories cast in the form of probabilistic generalizations, providing policy-makers a means of diagnosing the significance of specific situations.[6]

The nature of the phenomena being addressed dictate a focus on decisionmaking and the details of how crisis military operations are controlled. This, in turn, requires a research design in which a small number of cases are examined in detail using the method of structured focused comparison, rather than a research design using a large number of cases and statistical methods to identify significant variables. The purpose of structured comparison of a small number of cases is to reveal the different causal patterns that can occur for the phenomena, and the conditions under which each distinctive causal pattern occurs.[7]

The dependent variable is whether on not inadvertent escalation occurs in an international crisis. For the purposes of this study, inadvertent escalation will be defined as any increase in the level or scope of violence in a crisis that was not directly ordered by national leaders

---

[6]Alexander L. George and Richard Smoke, Deterrence in American Foreign Policy: Theory and Practice (New York: Columbia University Press, 1974), pp. 509-512; Alexander L. George, "Case Studies and Theory Development: The Method of Structured, Focused Comparison," in Paul Gordon, ed., Diplomacy: New Approaches in History, Theory, and Policy (New York: The Free Press, 1979), pp. 59-60.

[7]Ibid.

or anticipated by them as being the likely result of their orders. The specific phenomena to be explained in this study are the interaction of military forces in crises and the impact of such interactions on crisis stability.

Empirical research on the use of United States naval forces in crises will used to develop a set of contingent generalizations explaining three aspects of the theory: (a) the conditions under which crisis interactions become stratified and decoupled, (b) the conditions under which tensions between political and diplomatic objectives arise and affect crisis decisionmaking in particular ways, and (c) the conditions that prevent stratified escalation dynamics from occurring. The analysis will define discrete patterns of tactical-level crisis interaction, each associated with a particular causal pattern. Because the patterns of tactical-level interaction are arrived at empirically, the patterns identified in this study probably will not cover the universe of interaction patterns--additional patterns could well be identified through further empirical research.

The research design is divided into three phases. The first phase will be an examination of the mechanisms of delegated command, the nature of tactical-level military interactions, and the use of United States naval forces as a political instrument. These topics address some of the greatest weaknesses in the crisis management literature. The purpose of this phase of the research is to clarify

existing concepts and, when necessary, to present new concepts of crisis military operations before commencing the case studies. This first phase of the research design will encompass Chapters IV, V, and VI.

The second phase of the research design will consist of a structured focused comparison of four cases in which United States naval forces were employed in crises: the 1958 Taiwan Straits Crisis, the 1962 Cuban Missile Crisis, the 1967 Middle East War, and the 1973 Middle East War. The purpose of this phase will be to develop contingent generalizations on the theory of stratified interaction. This second phase of the research will be presented in Chapter VII.

To develop the contingent generalizations, eight questions addressing specific aspects of the theory will be answered through structured focused comparison. The first three questions address the conditions necessary for stratified interaction to occur: delegated control, tight coupling, and acute crisis.

Question 1. To what degree were interactions between the forces of the two sides at the scene of the crisis the result of actions taken in accordance with mechanisms of delegated control, rather than direct control by national leaders? If direct control was attempted, to what degree were national leaders able to exercise constant, real-time, positive control of operational decisions? If direct

control was nominally in effect but not in fact being exercised on a real-time basis, to what degree did on-scene commanders rely on guidance in mechanisms of delegated control relative to the direct guidance they received?

Question 2. Were the forces of the two sides at the scene of the crisis tightly coupled? Were on-scene (tactical level) commanders vertically integrated with sensors providing sufficient information on the adversary's on-scene forces to allow them to develop a picture of the adversary's moves and intentions independent of information provided to national leaders? Were tactical moves by each side quickly detected by the other side, prompting on-scene commanders to make (or request authorization to make) counter moves in order to preserve or improve their tactical situation?

Question 3. Were the forces of the two sides being used by their national leaders as a political instrument to convey deterrent or compellent military threats toward the other side in support of crisis bargaining? Were the forces of the two sides engaged in a test of capabilities under restrictive ground rules as a result of a challenge to a commitment being met by an effort to defeat that challenge without escalation? To what degree did interactions between the on-scene forces of the two sides influence the perceptions held by national leaders of the probability of war breaking out?

The fourth question addresses the first corollary to the theory of stratified interaction, that interactions can become decoupled in a crisis. There are seven potential causes of decoupling: communications and information flow problems, impairment of political-level decisionmaking, a fast-paced tactical environment, ambiguous or ambivalent orders, tactically inappropriate orders, inappropriate guidance in mechanisms of indirect control, and deliberate unauthorized actions by military commanders. More than one of these factors can occur simultaneously, further increasing the likelihood of unexpected escalation. The causes of decoupling are explained in Chapter III.

To establish that stratified interactions became decoupled in a crisis requires two findings: The first is that one or more of the seven seven potential causes of decoupling was present, creating opportunities for decoupling to occur. The second, and usually more difficult to establish, finding is that operational decisions made by tactical-level decisionmakers differed from the decisions that political-level decisionmakers probably would have made in order to coordinate those actions with their political-diplomatic strategy for resolving the crisis.

Question 4. Did interactions at the tactical and political levels become decoupled during the crisis? Did any of the potential causes of decoupling arise during the crisis? If conditions for decoupling existed, did national

leaders perceive the operational decisions made by the on-scene commander as interfering with or not supporting their political-diplomatic strategy for resolving the crisis? If momentary decoupling occurred in the crisis, was direct command immediately reimposed or did it initiate a decoupled interaction sequence?

The fifth question addresses the second corollary to the theory of stratified interaction, that the security dilemma is stratified in crises. The implication of this is that decisionmakers at the political and tactical levels can hold different perceptions of the offense-defense balance, vulnerability to preemption, and the need to strike first.

Question 5. Did national leaders and on-scene commanders hold different perceptions of the vulnerability of on-scene forces to preemption and the need to strike first in the event of an armed clash? Did actions taken with on-scene forces by national leaders for political signaling purposes generate tactical situations in which the on-scene commander perceived a vulnerability to preemption and a need to strike first should an armed clash erupt? Did actions taken for political purposes prompt the adversary's forces to take compensatory actions to reduce their vulnerability or to improve their ability to strike first?

The sixth question addresses the third corollary to the theory of stratified interaction, that escalation dynamics can be stratified in a crisis. A limitation

imposed on this study by the circumstances of post-World War II history is that no crises during the period escalated to a war in which the United States was a participant.[8] The absence of cases resulting in war precludes using the outcomes of the crises, in the sense of whether or not war occurred and the manner in which crises escalate to war, as dependent variables. Thus, the research design cannot address what would otherwise be the most interesting question, the circumstances under which decoupled, stratified interactions generate stratified escalation dynamics leading uncontrollably to war. Although this question cannot be addressed directly, research will be done to identify conditions which may have inhibited stratified escalation dynamics from occurring.

Question 6. When stratified interactions become decoupled, what factors inhibit escalation dynamics from occurring at the tactical level? When tactical-level interactions do begin escalating, what factors inhibit escalation dynamics from being transmitted upward to the

---

[8] I exclude the 1964 Tonkin Gulf Incident as a crisis that escalated to war because the incident did not lead to immediate and sustained U.S. intervention in the war. Although the U.S. commenced bombing North Vietnam after the incident, the bombings were in retaliation for subsequent attacks on U.S. forces in the South. Significant escalation of the U.S. role in the war, in the form of ground combat troops, did not occur until seven months after the Tonkin Gulf Incident. The decisions to escalate the U.S. role were made after months of deliberation, not under conditions of crisis as defined in this study.

strategic and political levels of interaction? Under what circumstances could these escalation-inhibiting factors break down, allowing a crisis to escalate uncontrollably to war?

The seventh question addresses the crisis management problems that arise when military forces are used as a political instrument in crises: the misperception dilemma and inadvertent military incidents.

Question 7. Did actions taken with military forces send inadvertent signals of hostility or acquiescence to adversaries, or inadvertent signals of encouragement or retrenchment to allies and friends? Were national leaders aware of the possibility of their military actions being misperceived and did this affect their decisionmaking? Did inadvertent military incidents occur and how did they affect efforts to manage the crisis? Under what circumstances did the inadvertent incidents occur and what factors contributed to their occurrence? Were national leaders aware of the possibility of inadvertent incidents and did this affect their decisionmaking?

The eighth question addresses the three tensions between political and military considerations that arise when military forces are used as a political instrument in crises: tension between political considerations and the needs of diplomatic bargaining, on the one hand, and military considerations and the needs of military operations, on

the other; tension between the need for top-level control of military options in a crisis, and the need for tactical flexibility and instantaneous decisionmaking at the scene of the crisis; and tension between performance of crisis political missions and readiness to perform wartime combat missions. Although the first source of tension (political versus military considerations) tends to pit military men against their civilian superiors, these tensions are not simply issues of civil-military relations. The second source of tension (level of operational control) can generate disputes between military commanders at the political, strategic, and tactical levels. The third source of tension (crisis missions versus readiness for wartime missions) involves significant conflicts between crisis military objectives and wartime military objectives, as well as conflicts between crisis political objectives and wartime military objectives. But all three tensions arise from the requirements of crisis management, the essence of which is placing political constraints on military operations.

Question 8. Did tensions arise between political considerations and the needs of diplomatic bargaining, on the one hand, and military considerations and the needs of military operations, on the other? Did tensions arise between the need for direct, positive, top-level control of military operations, and the need for tactical flexibility and instantaneous decisionmaking at the scene of the

crisis? Did tensions arise between performance of crisis missions and maintaining or increasing readiness to perform wartime missions? If any of these three tensions arose, how did they affect political-level and tactical-level decisionmaking? Are such tensions related to decoupling of stratified interactions and the occurrence of stratified escalation dynamics?

The third phase of the research design will consist of a structured, focused comparison of four cases in which a U.S. Navy ship was attacked during peacetime or crisis operations: the 1964 Tonkin Gulf Incidents, the 1967 Israeli attack on the USS Liberty, the 1968 North Korean seizure of the USS Pueblo, and the 1987 Iraqi attack on the USS Stark. Peacetime attacks on Navy ships are a particular concern due to their escalatory potential (which is discussed in the following section). The purpose of this phase will be to further develop and refine contingent generalizations on the theory of stratified interaction. The focus will be on how the naval and military chain of command reacted to the attack and whether or not crisis management problems arose from that reaction. The third phase will be presented in Chapter VIII.

To further develop the contingent generalizations, four of the previous eight questions will again be answered in a structured, focused comparison. The four questions address decoupling of stratified interactions, stratified

escalation dynamics, misperceptions, and political-military tensions.

Question 1. Did interactions at the tactical and political levels become decoupled during or after the attack on the Navy ship? Did conditions for decoupling arise during the crisis? If conditions for decoupling existed, did national leaders perceive the operational decisions made by the on-scene commander as interfering with or not supporting their political-diplomatic strategy for dealing with the attack? If momentary decoupling occurred in the crisis, was direct command immediately reimposed or did it initiate a decoupled interaction sequence?

Question 2. When stratified interactions become decoupled, what factors inhibit escalation dynamics from occurring at the tactical level? When tactical-level interactions do begin escalating in violence, what factors inhibit escalation dynamics from being transmitted upward to the strategic and political levels of interaction? Under what circumstances could these escalation-inhibiting factors fail, allowing a crisis to escalate uncontrollably to war?

Question 3. Did actions taken with military forces send inadvertent signals of hostility or acquiescence to adversaries, or inadvertent signals of encouragement or retrenchment to allies and friends? Were national leaders aware of the possibility of their military actions being misperceived and did this affect their decisionmaking? Did

inadvertent military incidents occur and how did they affect efforts to manage the crisis? Under what circumstances did the inadvertent incidents occur and what factors contributed to their occurrence? Were national leaders aware of the possibility of inadvertent incidents and did this affect their decisionmaking?

Question 4. Did tensions arise between political considerations and the needs of diplomatic bargaining, on the one hand, and military considerations and the needs of military operations, on the other? Did tensions arise between the need for direct, positive, top-level control of military operations, and the need for tactical flexibility and instantaneous decisionmaking at the scene of the crisis? Did tensions arise between performance of crisis missions and maintaining or increasing readiness to perform wartime missions? If any of these three tensions arose, how did they affect political-level and tactical-level decision-making? Are such tensions related to decoupling of stratified interactions and the occurrence of stratified escalation dynamics?

Focus on Naval Forces

As was noted in the research design, the cases to be examined all concern the use of United States naval forces as a political instrument in crises and peacetime attacks on U.S. Navy ships. There are four reasons for this. First,

of the branches of the U.S. armed forces, the Navy is the
service called upon most often to respond to crises. The
Navy is on the cutting edge of crisis management. Second,
American leaders and many analysts perceive naval forces as
having important advantages over other types of forces for
crisis response. Third, in spite of the frequency of use
and perceived advantages of naval forces, some U.S. Navy
officers and civilian analysts feel that the role of naval
forces as a political instrument is not well understood.
Fourth, some analysts believe that naval forces have a
greater escalatory potential than do other forces.

Naval forces have long had an important role in the
foreign policies of maritime nations. The United States
Navy in particular has often been called on to serve as an
instrument of national policy. Data on the employment of
the U.S. armed forces as a political instrument collected by
Barry M. Blechman and Stephen S. Kaplan show that U.S. Navy
units were employed in 177 of 215 incidents (83%) between
1945 and 1975, while a follow-on study by Philip D. Zelikow
found that U.S. Navy units were employed in 31 of 44
incidents (70%) between 1975 and 1982.[9] The U.S. Navy
supports peacetime foreign policy objectives through a
variety of missions, ranging from routine port visits and

---

[9]Barry M. Blechman and Stephen S. Kaplan, _Force
Without War_ (Washington, DC: Brookings Institution, 1978),
pp. 38-44; Philip D. Zelikow, "Force Without War, 1975-82,"
Journal of Strategic Studies 7 (March 1984): 29-54.

"showing the flag," to presence in strength at the scene of a conflict and retaliatory attacks against hostile nations. Wartime combat missions are the fundamental raison d'etre of navies, but peacetime political missions are their most common employment.

The U.S. Navy is the branch of the armed forces most commonly employed as a political instrument due to naval forces being perceived as having several inherent advantages for that role. The greatest advantage of naval forces stems from the medium in which they operate: naval vessels are free to roam the high seas (the oceans outside of territorial waters) without restrictions, asserting freedom of the seas--a principle well-established in international law.[10] The ability of naval forces to establish a visible U.S. presence in international waters near the scene of a crisis without intruding into disputed territory or immediate need of politically sensitive shore bases is an advantage not shared by land-based forces. The oceans provide naval forces with wide geographic reach, only the few nations without sea coasts and beyond the reach of carier aircraft are not readily influenced by sea power.

---

[10]Laurence W. Martin, The Sea in Modern Strategy (New York: Praeger, 1967), p. 67; Ken Booth, Navies and Foreign Policy (New York: Crane Russak, 1977), pp. 33-35; Hedley Bull, "Sea Power and Political Influence," in Jonathan Alford, ed., Sea Power and Influence: Old Issues and New Challenges (London: International Institute for Strategic Studies, 1980), p. 8.

The mobility and flexibility of naval forces are assets highly valued by national leaders. Naval forces are readily moved to a tension area, maneuvered to signal intentions and resolve, and withdrawn when U.S. objectives are achieved. Endurance, the ability to remain on station in a tension area for a prolonged period of time, is another important attribute of naval forces. The endurance of naval forces allows national leaders to send Navy ships to a tension area and then wait and see what develops. Although naval forces in a presence role serve primarily as a visible symbol of U.S. power and influence, their combat strength is a central element in their role. The ability of naval forces to project power ashore on short notice with naval gunfire, carrier airpower, cruise missiles, and Marine troops provides national leaders with a wide range of military options for conveying carefully crafted threats in support of diplomatic bargaining. Equally important, these combat capabilities also provide options for seeking a military solution to the crisis should it become necessary.[11]

---

[11] James Cable, Gunboat Diplomacy, 1919-1979, Second Edition (New York: St. Martin's Press, 1981), p. 67; Edward N. Luttwak, The Political Uses of Sea Power (Baltimore: Johns Hopkins University Press, 1974), p. 1. For an illustration of the range of options that naval forces provide the President in a crisis, see Floyd D. Kennedy, Jr.'s description of the 1958 Taiwan Straits Crisis in "The Creation of the Cold War Navy, 1953-1962," in Kenneth J. Hagan, ed., In Peace and War: Interpretations of American Naval History, 1775-1984, Second Edition (Westport, CT: Greenwood Press, 1984), pp. 316-317.

In contrast, land-based air and ground forces face numerous political, legal, and logistical constraints on their ability to be inserted into a tense area. They often require prepared bases (at least runways), and may not be welcome on foreign soil. Nationalism is a powerful emotion in many countries, particularly former colonies, and even nations desiring U.S. support may be hesitant to incur the domestic political strife that a foreign military presence can ignite. Land-based forces have a long and heavy logistical tail that makes them a cumbersome political instrument--they cannot be rapidly deployed other than in small units with low endurance, and once inserted can be difficult to withdraw. Deployment of land-based forces by air, or even use of long-range bombers for a show of force, can be precluded by reluctant allies and other nations refusing passage through their air space or refusing landing rights to refuel.

Employment of land-based forces normally entails inherently greater risks than employment of naval forces due to the much stronger political signals sent by forces ashore and their vulnerability to a wider range of threats. Because land-based forces imply a greater degree of permanence than do naval forces, land-based forces can signal a stronger and less flexible of commitment. Even if a strong signal of commitment was intended, the fact that land-based forces are difficult to move can inadvertently

create an actual degree of commitment greater than had been intended.[12]

Observers of naval diplomacy have concluded that changes in the structure and conduct of international politics since the end of World War II have been the primary factors causing maritime powers, particularly the United States, to place greater emphasis on the use of naval forces as a political instrument relative to land-based air and ground forces. Starting from the perspective of Robert E. Osgood and Robert W. Tucker that the destructiveness of nuclear war and the danger of conflicts escalating to nuclear war impose constraints on and "regulate" the use of force,[13] James A. Nathan and James K. Oliver contend that the superpowers have had to search for usable and controllable forms of military power--instruments of force which are both potent and responsive to the need for limits

---

[12]Martin, pp. 143, 146; Cable, p. 67; Luttwak, p. 1. The superiority of naval forces over land-based troops and aircraft was clearly demonstrated during the 1958 Lebanon Crisis, when Marines were landed with carrier air cover exactly when the President specified with only twelve hours notice. Severe logistical problems delayed the deployment of Air Force and Army units to the theater. See Kennedy, pp. 320-322.

[13]Robert E. Osgood and Robert W. Tucker, Force, Order and Justice (Baltimore: Johns Hopkins University Press, 1967), pp. 2-40. U.S. naval officers, in somewhat less analytic terms, had been asserting the existence of a "regulated" strategic environment for some time. See Admiral Arleigh A. Burke, "The U.S. Navy's Role in General War and Conflict Short of General War," Naval War College Review 11 (April 1959): 7-11.

on their use. They conclude that naval power has been the type of force best suited for use under these constraints, largely due to the advantages described above.[14] Similarly, James Cable has observed that "some of the constraints on the use of American military power to exert international influence are also such as almost to encourage reliance on limited naval force for this purpose."[15] Other observers have suggested that domestic political constraints in the United states have also caused naval forces to be favored over the other armed forces.[16] Thus, there is reason to believe that in the future naval forces will continue to be

---

[14]James A. Nathan and James K. Oliver, The Future of United States Naval Power (Bloomington: University of Indiana Press, 1979), pp. 17-18, 35. Also see Burke, pp. 9-11; and Rear Admiral John D. Chase, "The Function of the Navy," U.S. Naval Institute Proceedings 95 (October 1969): 30-32. Chase's description of Pax Ballistica is essentially the same as the regulated strategic environment described by Osgood and Tucker, but preceded it by a decade.

[15]Cable, p. 28.

[16]Commander Dennis R. Neutze, a Navy lawyer, has suggested that the 1973 War Powers Act, which requires the President to consult with Congress when U.S. forces are introduced into hostilities or a situation of imminent involvement in hostilities, makes naval forces preferable to land-based forces. Because naval forces can be deployed near the scene of a conflict without actually being introduced into hostilities, they do not activate the consultation requirement or the sixty-day limit in the War Powers Act until hostilities are actually initiated. See Commander Dennis R. Neutze, "Bluejacket Diplomacy: A Juridical Examination of the Use of Naval Forces in Support of United States Foreign Policy," JAG Journal 32 (Summer 1982): 133-134. Although his argument has merit, Presidents have tended in practice to ignore the War Powers Act when compliance with it would have interfered with their policy objectives--much to the displeasure of Congress.

the branch of the armed forces favored by United States leaders for crisis response.

The role of naval forces as a political instrument in peacetime received a great deal of attention in the U.S. Navy in the early 1970s. Despite the efforts made to develop concepts and principles of "naval presence," as peacetime naval employment was known, there remains dissatisfaction with our understanding of such political missions. Admiral Stansfield Turner stated in 1977, "I think that we who exercise naval presence do not know enough about how to fit the action to the situation: how to be sure that the force we bring to bear, when told to help in some situation, is in fact the one most appropriate to the circumstances."[17] Civilian analysts have echoed his concern. In a discussion of the relative importance of peacetime naval missions, Geoffrey Till emphasized that assessment of such issues "requires a full and proper understanding what naval diplomacy is, exactly how it works and what its requirements are." But he goes on to warn that "formulations of the strategy of naval diplomacy have as yet some way to go before these things are achieved."[18] When naval officers

---

[17]Stansfield Turner, "Designing a Modern Navy: A Workshop Discussion," in "Power at Sea: II. Super-Powers and Navies," Adelphi Papers No. 123 (London: International Institute for Strategic Studies, 1976), p. 28.

[18]Geoffrey Till, Maritime Strategy and the Nuclear Age, Second Edition (New York: St. Martin's Press, 1984), p. 214.

and analysts alike express concern that the use of United
States naval forces as a political instrument is not well
understood, there are grounds for questioning how well
United States leaders understand the implications of
employing naval forces for crisis response.

Several observers have expressed concern over the
escalatory dangers associated with the employment of naval
forces. Of particular concern to some observers is the
escalatory pressure that can arise when a U.S. Navy ship is
attacked. Former White House aide Chester Cooper,
commenting on the strong Senate reaction to the 1964 Tonkin
Gulf Incident, described the emotions aroused by attacks on
United States ships:

> There is something very magical about an attack on an
> American ship on the high seas. An attack on a
> military base or an Army convoy doesn't stir up that
> kind of emotion. An attack on an American ship on the
> high seas is bound to set off skyrockets and the 'Star
> Spangled Banner' and 'Hail to the chief' and
> everything else.[19]

George H. Quester and Sean M. Lynn-Jones have expanded upon
Cooper's remarks. Noting that "It is dreadfully dangerous
to sink a major power's warship today," Quester warns that
"the warships of the world have become highly prized invest-
ments, such that their loss would be likely to enrage the
publics and governments that matter back home--enrage them
enough to trigger off escalations that neither side might

---

[19]"The 'Phantom Battle' that Led to War," U.S. News
and World Report, July 23, 1984, p. 66.

have wanted, thus setting up the deterrence and bluff mechanisms that are at the heart of 'chicken'."[20] Along the same lines, Lynn-Jones observed that "Under conditions of international tension and superpower rivalry, public opinion in a liberal democracy is likely to demand retaliation after a provocation by a major rival. Naval incidents seem to elicit particularly emotional responses in the United States." He goes on to add that "Is is, of course, relatively unlikely that a naval incident could provoke a nuclear exchange between the United States and the Soviet Union. ...An incident could, however, increase tensions and needlessly disrupt negotiations or other political discourse, much as the U-2 incident of 1960 forced the cancellation of the Khrushchev-Eisenhower summit."[21]

Another view is that there is a greater risk of nuclear war erupting at sea than ashore. This argument has been made forcefully by Desmond Ball:

> The possibility of nuclear war at sea must be regarded as at least as likely as the occurrence of nuclear war in other theaters. Indeed, there is probably a greater likelihood of accidental or unauthorized launch of sea-based nuclear weapons, and the constraints on the authorized release of nuclear weapons are possibly more relaxed than those that pertain to land-based systems. Further, there are

---

[20] George H. Quester, "Naval Armaments: The Past as Prologue," in George H. Quester, ed., Navies and Arms Control (New York: Praeger Publishers, 1980), pp. 6-7.

[21] Sean M. Lynn-Jones, "A Quiet Success for Arms Control: Preventing Incidents at Sea," International Security 9 (Spring 1985): 164.

several important factors that make it likely that any
major conflict at sea would escalate to a strategic
nuclear exchange relatively quickly.[22]

Incidents at sea between American and Soviet forces have
been identified as a potential catalyst for the nuclear
escalation dangers described by Ball. As John Borawski
notes: "The 1967 Israeli sinking [sic] of the USS Liberty,
and the subsequent US uncertainty as to whether a Soviet
ship had attacked the Liberty, is often cited as an example
of the type of nuclear Sarajevo that could inadvertently
lead to war."[23] Thus, there are at least prima facie
reasons for concern that the use of naval forces as a
political instrument in crises has an escalatory potential

---

[22]Desmond Ball, "Nuclear War at Sea," International
Security 10 (Fall 1985): 28-29. The factors Ball identifies
are the occurrence of accidents at sea, the attractiveness
of ships as nuclear targets, the nuclear weapons launch
autonomy of naval commanders, dual-capable weapons systems
and platforms, offensive Navy anti-submarine warfare (ASW)
strategy (including attacks on Soviet strategic ballistic
missile submarines), incentives for Soviet preemption
arising from the vulnerability of Navy ASW and command and
control systems, the Navy doctrine of offensive operations
in forward areas, Navy tactical nuclear weapons doctrine,
Soviet doctrine for war at sea, and lack of Navy contingency
planning for limiting escalation in a war at sea. Also see
Barry R. Posen, "Inadvertent Nuclear War? Escalation and
NATO's Northern Flank," International Security 7 (Fall
1982): 28-54; Eric J. Grove, "The Maritime Strategy and
Crisis Stability," Naval Forces 8 (6/1987): 34-44.

[23]John Borawski, "Risk Reduction at Sea: Naval
Confidence-Building Measures," Naval Forces 3 (1/1987):
18. It must be noted that Liberty was not sunk in the
attack. As will be discussed in Chapter VII, no U.S. Navy
commander in the chain of command thought that the Soviets
had conducted the attack, and the commanders in the
Mediterranean knew for a fact that the Soviets could not
have conducted the attack.

that has not been adequately addressed in studies of naval diplomacy and crisis management.

The characteristics of naval forces that give them their advantages as a political instrument also generate three serious potential problems for crisis management. First, the political signals sent by naval forces are particularly prone to being misperceived, inadvertently sending the wrong signals to allies and adversaries. Second, the nature of the maritime environment, in which forces of the two sides in a crisis routinely operate at point blank range, exacerbates problems of maintaining control of events. Third, the nature of the naval warfare environment, which places a premium striking first in tactical engagements, exacerbates problems of crisis stability and escalation control. These problems of using naval forces as a political instrument are examined in greater detail in Chapter VI.

To summarize, U.S. Navy crisis operations and peacetime attacks on U.S. Navy ships will be used as the historical cases for this study because the Navy is the service called upon most often to respond to crises, American leaders and many analysts perceive naval forces as having important advantages over other types of forces for crisis response, some U.S. Navy officers and civilian analysts feel that the role of naval forces as a political instrument is not well understood, and some analysts believe

that naval forces have a greater escalatory potential than
do other forces.

Cases and Case Selection

Two sets of historical cases will be used as sources
of empirical data, one set of cases for each phase of the
research design. These cases will be used as sources of
empirical data for deriving the contingent generalizations.
Although essentially the same questions will be asked in
each case, full-scale case studies will not be conducted.

Empirical data for the second phase of the study will
come from four cases in which United States naval forces
were employed in crises: the 1958 Taiwan Straits Crisis, the
1962 Cuban Missile Crisis, the 1967 Middle East War, and the
1973 Middle East War. The criteria for case selection in
the first phase of the research were (a) significant U.S.
naval operations were conducted which influenced the outcome
of the crisis, (b) naval operations were conducted in the
immediate proximity of adversary naval forces or land-based
forces that could threaten naval forces, and (c) there was a
possibility of fighting erupting between the United States
and the other side in the crisis.

Given the large number of crises in which the U.S.
navy has played an important role, case selection was
particularly difficult for this phase of the study. Among
the more prominent cases considered and rejected were the

1954 Quemoy-Matsu Crisis, the 1958 Lebanon Crisis, the 1970

Jordanian Crisis, and the 1971 Indo-Pakistani War. Although

the 1954 Quemoy-Matsu Crisis and the evacuation of the

Tachen Islands were perhaps as serious as the 1958 case,

there was less tactical-level interaction because China

ceased its harassment of the islands while the U.S. Navy was

on the scene (thus making a naval confrontation an unlikely

source of escalation). The Navy role in the 1956 Suez

Crisis was limited to evacuation of civilians, there was

little tactical-level interaction, and little concern that

the crisis would escalate to war. There was little tactical-

level interaction in the 1958 Lebanon Crisis, little concern

that the crisis would escalate to war with the Soviet Union,

and, after the Marines were landed, little concern that the

U.S. would be involved in a civil war. There was minor

tactical-level interaction in the 1970 Jordanian Crisis, but

the Navy role was small and there was little concern that

the crisis would escalate to war. Although there was

tactical-level interaction and concern among Navy officers

over the Soviet naval threat in the 1971 Indo-Pakistani War,

the Navy role was limited and there was little concern that

the crisis would escalate to war.

Empirical data for the third phase of the study will

come from four cases in which U.S. Navy ships were attacked

in peacetime: the 1964 Tonkin Gulf Incident, the 1967

Israeli attack on the USS Liberty, the 1968 North Korean

seizure of the USS Pueblo, and the 1987 Iraqi attack on the USS Stark. The criteria for case selection were (a) the attack was on a U.S. naval vessel, and (b) the attack occurred during a crisis or under circumstances that could have provoked a U.S. military response. The second criterion excludes limited war situations, such as the Korean War and the Vietnam War. The four cases that were selected are of interest because they come closest to illustrating the circumstances in which stratified interactions could become decoupled and stratified escalation dynamics occur.

Although this study will focus on the use of naval forces as a political instrument, the limitations of this approach are recognized. Non-naval activities, particularly diplomatic efforts, may be as important to the success of naval diplomacy as the actions of the ships at the scene of tensions. As Ken Booth has reminded, "naval diplomacy is a matter of diplomats on land as well as ships at sea and of the role the former can play to ensure that naval messages are not misperceived."[24] Furthermore, it can be difficult to separate the particular contribution of naval force from the overall diplomatic and military effort made to resolve a dispute. Commander James F. McNulty has observed that "it

---

[24]Ken Booth, "Foreign Policies at Risk: Some Problems of Managing Naval Power," Naval War College Review 29 (Summer 1976): 15.

is usually difficult to assess the effect of Naval Presence alone on decisions which of necessity are made as the outcome of reaction to a broad range of American signals-- military, economic, and political--perceived by other nations."[25] Thus, caution must be exercised when attempting to assess the role of naval forces in achieving a given political outcome.

Organization of the Study

This study will begin, in Chapter II with a review and critique of the literature on crises and crisis management. Chapter III defines the theory of stratified interaction and its corollaries. Chapter IV examines the mechanisms of indirect control, providing background on the command and control procedures that influence the stratification of crisis interactions. Chapter V discusses tactical-level military interaction, illustrating the range of interactions that can occur in crises. Chapter VI explores the use of naval force as a political instrument in crises, applying the theory of stratified interaction under the particular operational circumstances surrounding the use of naval forces. Examining crisis naval operations will also reveal the political-military tensions that arise when crisis

---

[25]Commander James F. McNulty, "Naval Presence -- The Misunderstood Mission," Naval War College Review 27 (September-October 1974): 28.

management objectives and military objectives are pursued simultaneously.

The next two chapters present the case studies. Chapter VII presents the second phase of the empirical research, the four case studies on crisis naval operations. Chapter VIII presents the third phase of the empirical research, the four case studies on peacetime attacks on navy ships. Chapter IX presents the findings of the case studies and presents the contingent generalizations on stratified interaction. Internal and external factors that appear to prevent stratified interactions from decoupling will also be discussed. Chapter X presents conclusions on the theory and suggests the policy implications of the findings.

CHAPTER II

USE OF FORCE IN CRISES:

A REVIEW AND CRITIQUE

Current concepts and theories on crises and crisis management have three serious weaknesses. First, the various political and military interactions that occur between the two sides in a crisis are assessed in the context of an implicit single interaction sequence model of crises. Second, the frequently observed phenomenon of United States leaders exercising close control over military operations in crises, combined with a lack of familiarity with military command and control procedures, has produced an erroneous view of the manner in which military forces are controlled in crises. Third, and derived from the two previous weaknesses, the concept of crisis stability is inadequately developed and there is a poor understanding of the escalation processes that could cause a crisis to escalate to war.

This chapter will review and critique the literature on crises and crisis management, developing in detail three major weaknesses in current concepts and theories. The first section will review basic concepts on international

crises, presenting the perspective on crises that will be used in this study. The second section will critique the concept of crisis interaction and explain the weaknesses in the single interaction sequence model that implicitly underlies existing crisis theories. The third section will review basic concepts of crisis management, focusing on the measures required for national leaders to maintain control of events in crises. The fourth section will review and critique the concept of crisis stability, explaining the weaknesses in current conceptions of crisis stability and presenting a definition that more accurately reflects the nature of crisis interaction. The final section will review a serious problem in crisis management--misperception of intentions and resolve--and present concepts that more accurately describe the problems decisionmakers face in trying to avoid misperceptions when using force in crises.

International Crises

Through journalistic and political license the term "crisis" has been stretched to describe a wide range of phenomena. Essentially any problem for which national leaders do not have a ready solution can, at the whim of pundits or politicians, be labeled a crisis. The broad definition of crisis used in the vernacular lacks sufficient precision for this study because it covers far too wide a range of political situations.

The focus of this study is on a particular category of international crisis: the "acute" international crisis.[1] This type of crisis has been defined by Oran R. Young as "a process of interaction occurring at higher levels of perceived intensity than the ordinary flow of events and characterized by: a sharp break from the ordinary flow of politics; a rise in the perceived prospects that violence will break out; and significant implications for the stability of some system or subsystem (or pattern of relationships) in international politics."[2] Phil Williams defines such a crisis as "a confrontation of two or more states, usually occupying a short time period, in which the probability of an outbreak of war between the participants is perceived to increase significantly."[3] Along these same lines, Glenn H. Snyder and Paul Diesing define a crisis as "a sequence of interactions between the governments of two or more sovereign states in severe conflict, short of actual war, but involving the perception of a dangerously high probability of war."[4] Finally, the definition of crisis

---

[1]Charles A. McClelland, "The Acute International Crisis," World Politics 14 (October 1961): 182-205.

[2]Oran R. Young, The Politics of Force (Princeton, NJ: Princetion University Press, 1968), p. 15.

[3]Phil Williams, Crisis Management (New York: John Wiley and Sons, 1976), p. 25.

[4]Glenn H. Snyder and Paul Diesing, Conflict Among Nations (Princeton, NJ: Princeton University Press, 1977), p. 6.

used in International Crisis Behavior Project headed by

Michael Brecher is that:

> a crisis is a situation with three necessary and
> sufficient conditions, deriving from a change in its
> external or internal environment. All three are
> perceptions held by the highest level decision-makers:
> 1. threat to basic values, with a simultaneous or
> subsequent
> 2. high probability of involvement in military
> hostilities, and the awareness of
> 3. finite time for response to the external value
> threat.[5]

Thus, the essential features of acute international crises

are a confrontation, short of war, between two sovereign

states, and a perception by national leaders of a

significantly increased danger of war breaking out, or at

least greatly increased uncertainty that war can be avoided.

Shortness of duration is has been used by some

observers to distinguish an acute crisis from other crises.[6]

This is normally done because national leaders tend to feel

severe time constraints and an urgent need to take immediate

action in crises, and because crises that drag out for weeks

or months lose their intense sense of danger as implicit

norms of behavior are tacitly established through actual

practice. However, prolonged crises are at least as likely

to occur as are acute crises, can have as great a potential

to escalate to war, and can be just as threatening to the

---

[5]Michael Brecher, Decisions in Crisis: Israel, 1967
and 1973 (Berkeley: University of california Press, 1980),
p. 1 (emphasis is his).

[6]Young, p. 15; Williams, p. 25.

national interests of the parties involved. Prolonged
crises are certainly worthy of investigation, and have been
somewhat neglected in the crisis literature due to the focus
on short-duration crises.[7] Additionally, an acute crisis
can arise during a prolonged crisis if either side takes an
action seriously violating the tacit norms of behavior being
observed. Although the phenomena of interest in this study
are most prominent in short-duration acute crises, they also
occur during the periods of acute crisis that can arise
during a prolonged crisis. Thus, while prolonged crises are
not excluded from this study, when they are addressed
attention will be focused on the periods of acute crisis
within them.

A feature of some crises, which can contribute to
crises being of short duration, is that national leaders
perceive themselves as acting under time constraints--action
must be taken immediately to avert unacceptable losses to
vital national interests.[8] The perception of time con-
straints held by leaders of one nation is usually induced or

_____

[7]Eliot A. Cohen, "Why We Should Stop Studying the
Cuban Missile Crisis," The National Interest No. 2 (Winter
1985/6): 6.

[8]Ole R. Holsti, Crisis Escalation War (Montreal:
McGill-Queen's University Press, 1972); p. 9; Charles F.
Hermann, "Some Issues in the Study of International Crisis,"
in Charles F. Hermann, ed., International Crises: Insights
from Behavioral Research (New York: The Free Press, 1972),
p. 13; Richard N. Lebow, Between Peace and War: The Nature
of International Crises (Baltimore: Johns Hopkins University
Press, 1981), p. 12; Brecher, p. 1.

exacerbated by the actions of the other side in the crisis, particularly if the crisis was provoked by an attempt at a *fait accompli*, or if an ultimatum was presented. The effect of a perception of time constraints is to raise the level of stress experienced by national leaders, possibly reducing the effectivenesss of their analysis and decisionmaking. Perception of time constraints is a variable rather than a parameter in crises, and can vary widely in intensity. In a prolonged crisis all the features of crisis are present except the perception of time constraints. Perception of time constraints tends to be strong in the type of crisis of interest to this study, but is not a necessary condition for an international crisis to exist.

Another feature of international crises is that national leaders perceive important national interests to be at stake in the conflict.[9] Such perceptions are particularly intense in acute crises. Examples of national interests commonly perceived as important enough to warrant a crisis include the security of the nation and its allies, spheres of influence or positions of regional political prominence, international principles such as freedom of the seas or rights of neutrals, and sources of strategic minerals or foodstuffs. The nation's reputation as a world power and its bargaining reputation have sometimes been included as

---

[9]Holsti, p. 9; Hermann, p. 13; Williams, p. 25; Lebow, **Between Peace and War**, p. 10.

interests which when threatened can provoke a crisis, but such interests normally become involved when a threat to one of the more concrete interests listed above arises, thus compounding the importance of the interest at stake. It is the threat to important national interests that generates the sense of urgency and perceived danger of war in crises.[10]

A crisis is fundamentally a bargaining relationship between the two sides.[11] Bargaining relationships are marked by interdependence: the ability of each side to achieve its objectives depends on the decisions and actions of both sides. Thus each side, in planning its own course of action, must take into account the objectives and anticipated course of action of the other side.[12] In crisis

---

[10]This definition of acute crisis omits surprise as a characteristic, which is included by Holsti, p. 10, and Hermann, p. 13. However, a crisis need not surprise national leaders by its occurrence, and could build up gradually from a prolonged dispute, so long as it arises in such a way as to give national leaders the perception that it threatens serious damage to important national interests. See Snyder and Diesing, p. 17.

[11]Thomas C.Schelling, The Strategy of Conflict (Cambridge, Mass.: Harvard University Press, 1960), pp. 187-203; Glenn H. Snyder, "Crisis Bargaining," in Charles F. Hermann, ed., International Crises: Insights from Behavioral Research (New York: The Free Press, 1972), pp. 217-256.

[12]A decision by one side to seek a military solution to a crisis, as in a fait accompli, does not necessarily eliminate bargaining as a feature of the crisis. The outcome still depends of the decision by the other side whether to resist or to sacrifice its interests in order to avoid war. Furthermore, bargaining may continue to achieve a final resolution of the dispute.

bargaining the two sides have common or complementary interests, as well as conflicting interests, otherwise one or both sides would opt for war rather than engage in crisis bargaining.[13] The advantage of viewing international crises as a bargaining relationship is that it highlights their fundamental political nature, which can be obscured by the military actions taken during crises. To emphasize bargaining is not to deny that crises are an intense form of strategic competition over interests perceived as being vital by national leaders on the two sides. Rather, the focus on bargaining provides a useful means for conceptualizing how strategic competition is conducted in crises.

Bargaining of some sort is, of course, present across the entire spectrum of international intercourse, from routine peacetime negotiations to full-scale war. But international crises stand apart from both peacetime diplomatic disputes and wartime military conflict due to their unique political-military nature. In crisis bargaining, varying combinations and sequences of persuasion, coercion and/or accommodation are applied in an effort at resolving the conflict on favorable terms. Although the threat of resort to force, even if only as a latent coercive

---

[13]Coral Bell refers to this as an "adverse partnership," which is marked by "solid common interests as well as sharp conflicting interests." The Conventions of Crisis (London: Oxford University Press, 1971), p. 50.

threat, is rarely ever totally excluded in peacetime diplomacy, and political accommodation is rarely ever totally excluded in wartime hostilities, in crises both the political and military dimensions are prominent. In an acute crisis the confrontation has intensified to the point that coercion--direct, implied, or even latent, including standing deterrent threats as well as specific threats related to the crisis--begins to dominate the relationship and the grounds for accommodation begin to shrink toward little more than a mutual desire to avoid war. The prominence of coercion has led some observers to classify crises as an intermediate status of relations between peace and war, combining elements of both peacetime accommodation and wartime coercion.[14]

Crises consist of a series of bargaining interactions between the two sides. Bargaining interactions include formal negotiations, official diplomatic communications, informal communications via intermediaries or the media, and actions taken to convey political signals. Focusing on interaction highlights the interdependence between the two sides. Decisions made by each side reflect decisions made by the other side as well as their own objectives, and the ability of either side to achieve its objectives is dependent upon decisions made by the other side.

---

[14]Glenn H. Snyder, "Crisis Bargaining," pp. 218, 240; Snyder and Diesing, p. 10.

Although a particular type of international crisis--
the acute crisis--has been identified as the focus of this
study, that category still covers a broad range of crisis
phenomena. Not all types of crises are relevant to this
study. It will thus be useful to review the typologies of
crises that have been proposed in the crisis literature.
Coral Bell distinguishes between adversary crises, those
between nations regarding themselves as adversaries, and
intramural crises, those among allies or members of a
regional organization. Williams, and Snyder and Diesing,
draw a similar distinction, excluding intra-alliance crises
from their analyses.[15] This is a useful because intra-
alliance crises, though they may be acute politically and
involve explicit coercion, rarely entail risk of war. Thus,
the scope of this study will be limited to adversary crises.

More extensive typologies of crises have also been
proposed. Young proposed six types of crises, based on how
they are initiated: an attempt at a _fait accompli_, applying
coercive pressure as an indirect response to undesirable
actions, a military response to nonviolent provocation,
military probe provoking a military response, a military
invasion provoking military resistance, and mutual
intervention in political upheaval in a third country.[16]

---

[15]Bell, p. 7; Williams, p. 24; Snyder and Diesing,
p.7.

[16]Young, p. 22.

These categories are not particularly useful for analysis of crises, and serve mainly to illustrate the range of actions that can provoke a crisis.

A more useful approach is to distinguish among different motives for provoking a crisis. Snyder and Diesing distinguish among three types of crises: the "coercive bargaining type," a confrontation arising from a challenge met by resistance, the "war scare" or "security dilemma" type, arising from fear of imminent attack, and the "prelude or pretext to an intended attack" type, provoked to justify a preplanned military move. A hypothetical category, "accidental crises," is excluded by Snyder and Diesing for lack of empirical evidence that such a crisis has occurred.[17] Lebow has proposed a similar scheme of four types of crises: the "justification of hostility" crisis, used as a causus belli for war, the "spinoff" crisis, a deliberate hostile act toward a third country taken to further the prosecution of a war in progress, the "brinkmanship" crisis, a challenge to a known interest or commitment of another country in expectation that the other country will be compelled to back down rather than fight, and the "accidental" crisis, caused by an undesired and unsanctioned provocation.[18]

---

[17]Snyder and D. .sing, pp. 11-17.

[18]Lebow, Between Peace and War, pp. 23-97. Lebow states the 1905 Dogger Bank Incident was an accidental crisis. See his "Accidents and Crises: The Dogger Bank Affair," Naval War College Review 31 (Summer 1978): 66-75.

Contrasting these two schemes, Snyder and Diesing's coercive bargaining crisis is the same as Lebow's brinkmanship crisis, and Snyder and Diesing's pretext to attack crisis is the same as Lebow's justification of hostility crisis. Snyder and Diesing's arguments for treating a war scare as a separate category of crisis are not persuasive. War scare crises are better viewed as a form of coercive bargaining crisis in which the security dilemma has a major impact on crisis stability. Lebow's spinoff crisis also will not be addressed as separate category because it does not address the motives for the crisis so much as the circumstances in which it arose. All spinoff crises fall into either the coercive bargaining or pretext to attack categories.

Accidental crises, which were excluded by Snyder and Diesing but included by Lebow, have not been adequately addressed in the crisis literature. The role of military accidents in provoking or exacerbating crises has received attention, but as yet there are not adequate concepts for dealing with the effects of accidents. This study will not treat accidental crises as a separate category. Rather, inadvertent military incidents will be viewed as provoking or exacerbating one of the other two major categories of crises--coercive bargaining (brinkmanship) or pretext to attack (justification for war)--depending on how the two sides respond to the incident.

We are thus left with two major categories of crises: coercive bargaining (brinkmanship) and pretext to attack (justification for war). Of these two categories, the coercive bargaining or brinkmanship crisis is the type of interest in this study. While a pretext to attack or justification of hostility crisis is certainly an acute crisis, and could well entail intensive bargaining, the outcome is preordained to be war.[19]

Because the role of force as a political instrument in crises can vary significantly depending on the nature of the crisis, this study will distinguish between two categories of crises: direct and indirect. A direct crisis is one in which the United States is in direct confrontation with

---

[19]Analytical problems can arise when trying to distinguish between "pretext to attack" and "coercive bargaining" crises. That the outcome of a crisis was war is insufficient to establish that crisis as having been a pretext, it must also be shown that the nation which precipitated the crisis desired war to be the outcome regardless of the response by the other side. Complicating this analysis it the possibility of dual motives in a coercive bargaining crisis: if the target nation immediately capitulates to all demands, the initiator suspends his war plans, but if the target nation resists, the initiator launches war using the crisis as a pretext. The motives of the nation precipitating the crisis can also change during the crisis. A crisis provoked as a pretext for war could have a non-war outcome if the target nation were to offer much larger concessions than the initiator had expected to gain through coercion. Conversely, a crisis provoked for coercion could result in war if the target nation is unwilling to accept the initiator's demands and the initiator then decides to use the crisis as grounds for war. Thus, while its is useful to distinguish among crises on the basis of motives, the possibility of dual motives and changes in motives must be recognized.

another nation. The seizure of the USS <u>Pueblo</u> in 1968 generated a direct crisis between the United States and North Korea. The term <u>direct</u> <u>superpower</u> <u>crisis</u> will be used to describe a direct crisis in which the Soviet Union is the adversary. The 1962 Cuban missile crisis was a direct superpower crisis. An <u>indirect</u> <u>crisis</u> is one in which the United States is involved because it is supporting a friend or ally who is a direct participant. The 1958 Taiwan Straits crisis was an indirect crisis between the United States, supporting allies on Taiwan, and China. The term <u>indirect</u> <u>superpower</u> <u>crisis</u> will be used to describe an indirect crisis in which the United States and the Soviet Union are brought into confrontation by a conflict between their respective allies or clients. The 1967 and 1973 Middle East Wars generated indirect superpower crises. The term indirect is used to convey the sense that the outcome of the crisis, whether or not a Soviet-American war results, can be influenced by the decisions of the third parties as well as the decisions made by the two superpowers.[20]

---

[20]See Williams, pp. 130-34. What is referred to in this study as an "indirect superpower crisis" is described as a "limited local war" by Yaacov Bar-Simon-Tov. The difference is primarily one of perspective: this study focuses on the crisis between the superpowers brought on by the limited local war, whereas Bar-Simon-Tov's study focuses on the local war itself. Bar-Simon-Tov provides a superb analysis of bargaining relationships in this type of crisis in <u>The Israeli-Egyptian War of Attrition, 1969-1970</u> (New York: Columbia University Press, 1980), pp. 17-20.

In summary, this study will address acute international crises, which are characterized by a confrontation, short of war, between two sovereign states, and a perception by national leaders of a significantly increased danger of war breaking out. A crisis is a bargaining relationship between the two sides, and, as such, consists of a series of interactions between them. Additional features of acute international crises are that national leaders perceive important national interests to be at stake in the conflict and tend to perceive themselves as acting under severe time constraints. This study will be limited to adversary crises, excluding intra-alliance crises. Of the two major categories of crises--coercive bargaining (brinksmanship) and pretext to attack (justification for war)--the coercive bargaining crisis is the type that is of interest in this study. Finally, this study will distinguish between two categories of crises: direct, in which the United States is in direct confrontation with another nation, and indirect, in which the United States is involved through support of a friend or ally that is a direct participant.

Interaction in Crises

Previous studies of international crises have implicitly viewed the various political and military interactions that occur between the two sides as a single interaction sequence. This can be seen in the definitions

of crisis given above: Young describes a crisis as "a process of interaction," while Snyder and Diesing describe a crisis as "a sequence of interactions."[21] Because both of these analytical schemes focus primarily on top-level decisionmaking, with little attention to decisionmaking by military commanders at the scene of a crisis, the flow of events in a crisis is implicitly viewed as a single sequence of actions and reactions. This perspective on crisis interaction will be referred to as the single interaction sequence model.

The single interaction sequence model does not accurately describe the complexity of crisis interaction. What actually occurs in a crisis is multiple interaction sequences that only partially influence each other. In a crisis, national political leaders on the two sides are interacting through diplomatic communications and political signalling, national military leaders are interacting through the actions taken with their forces, and military forces in the field are interacting as they respond to orders from higher authorities and the actions of adversary forces. Such multiple interaction sequences, evolving simultaneously and semi-independently, arise when national leaders do not make all operational decisions themselves, but must delegate significant decisionmaking authority to

---

[21]Young, p. 15; Snyder and Diesing, p. 6.

subordinates. The single interaction sequence model views these multiple interaction sequences as a single flow of events.

The weakness of the single interaction sequence model is that in subtle ways it leads towards a unitary actor perspective of national behavior. In its pure form, the unitary actor model assumes that all actions taken by a nation are at least authorized, if not specifically ordered, by national leaders. The unitary actor model is typically used in strategic analyses of the national interests, objectives, and strategies that lead to crises. Countries are treated as entities having interests, objectives, and strategies. The role of organizations and individuals in the formulation and execution of policy are essentially ignored. Accidents and the possibility of national leaders losing control of the momemtum of military actions receive scant attention. The single interaction sequence model is compatible with the implicit, even indavertent, assumption that national leaders have authorized or are in direct control of the actions taken by their forces in a crisis.

The weaknesses in the unitary actor model are well recognized, which has lead to widespread use of the bureau- cratic politics and organizational process models for analysis of international crises. The bureaucratic politics model recognizes that the policy perspectives held by parti- cipants in decisionmaking are shaped by the organizations

they represent, and that the policy recommendations made by participants will be influenced by the parochial interests of their organizations. In the organzational process model governmental action is viewed as organizational output: the decisions of government leaders trigger organizational routines, which primarily determine the nature of the actions taken. Organizational activity consists largely of enactment of preestablished routines--the standard operating procedures and programs which constitute an organization's repertoire.[22]

Although the bureaucratic politics and organizational process models provide a more accurate description of decisionmaking than does the unitary actor model, they are not without their faults. The bureaucratic politics model tends to treat all policy recommendations made to the President and his closest advisors as having been motivated primarily by parochial bureaucratic self-interests. There is thus an inherent bias toward interpreting evidence of policy disagreements or actions not ordered by the President as evidence of bureaucratic politics.[23] The model does not

[22]Graham T. Allison, Essence of Decision: Explaining the Cuban Missile Crisis (Boston: Little, Brown and Co., 1971), pp. 78-96, 162-181.

[23]Stephen D. Krasner, "Are Bureaucracies Important? (Or Allison Wonderland)," Foreign Policy No. 7 (Summer 1972): 159-79; Robert J. Art, "Bureaucratic Politics and American Foreign Policy: A Critique," Policy Sciences 4 (December 1973): 467-90; Donald Hafner, "Bureaucratic Politics and 'Those Frigging Missiles': JFK, Cuba and U.S.

recognize two other possibilities: first, that conflicting policy recommendations may be based on considerations of national interest and the feasibility of various courses of action, rather than bureaucratic self-interest, and, second, that cabinet-level officals may well base recommendtions on personal policy preferences or political considerations, rather than on the interests of their bureaucracies.[24]

The bureaucratic politics model fails to recognize that national policies can be shaped by factors other than the interplay of bureaucratic politics. This is apparent in the "cult of the offensive" theory of the origins of World War I. According to this theory, a principle cause for the outbreak of war was that the armed forces of the European powers had a bias for offensive military doctrines. Their bias for the offensive is portrayed as being the result of parochial organizational interests--autonomy from civilian control, larger budgets, and prestige--as opposed to rational analysis of national strategic interests.[25]

Missiles in Turkey," Orbis 21 (Summer 1977): 307-33; Barton J. Bernstein, "The Cuban Missile Crisis: Trading the Jupiters in Turkey?" Political Science Quarterly 95 (Spring 1980): 103.

[24]Alexander L. George, Presidential Decsisionmaking in Foreign Policy (Boulder, CO: Westview Press, 1980), pp. 114-16.

[25]Steven Van Evera, "The Cult of the Offensive and the Origins of the First World War," International Security 9 (Summer 1984): 58-107; Jack Snyder, "Civil-Military Relations and the Cult of the Offensive, 1914 and 1984," International Security 9 (Summer 1984): 108-146.

The "cult of the offensive" theory has come under criticism for failing to recognize that factors other than the parochial interests of the militaries also drove the preference for offensive doctrines. In particular, some of the European powers needed offensive strategies to fulfill alliance commitments and others had definite policies of expansion and aggrandizement. Thus, in this example, the bureaucratic politics model resulted in explanatory factors other than parochial organizational interests being ignored and over-emphasis of the role of such interests in shaping national strategies. Additionally, Jack S. Levy criticizes the theory for its emphasis on bureaucratic routines as the causal link between crises and war. Levy argues that while bureaucratic factors may well lead to an offensive bias, that bias does not inevitably cause crises to escalate to war--additional, non-bureacratic, conditions must be present for war to erupt.[26]

The organizational process model has similar problems. It implicitly accepts the simple public administration distinction between policymaking and policy implementation. Once the President has decided on a course

_____

[26]Jack S. Levy, "Organizational Routines and the Causes of War," International Studies Quarterly 30 (June 1986): 193-222; and Scott D. Sagan, "1914 Revisited: Allies, Offense, and Instability," International Security 11 (Fall 1986): 151-75. Also see the exchange between Snyder and Sagan in "Correspondence," International Security 11 (Winter 1986-87): 187-98.

of action, government organizations serve only to carry out his orders--essentially devoid of their own policymaking authority. Organizational routines serve only to explain how presidential orders are corrupted in the process of implementation. This, in turn, leads to the implicit assumption that all actions taken by a nation during a crisis either are ordered by national leaders in pursuit of their policy objectives, or should not have occurred and therefore represent a loss of control over events.

This raises the second weakness in the crisis management literature, which is that it is based on an erroneous view of the manner in which military forces are controlled in crises. This apparently resulted from the frequently observed phenomenon of United States leaders exercising close control over military operations in crises, combined with a lack of familiarity with military command and control procedures. The crisis management literature typically describes the control of crisis military operations as being highly centralized, with top-level civilian authorities exercising direct control--in contrast to routine peacetime operations, which are described as highly decentralized and having little involvement of civilian political authorities. This description fails to grasp the complexity of military command and control.

Even in crises, military commanders are delegated significant authority to make operational decisions on the

employment of their forces--including specified decisions on the use of force. Under certain circumstances, spelled out when the delegation of authority was made, military commanders can use conventional weapons without seeking permission from higher authorities. The scope of their authority is spelled out in a variety of documents, which collectively will be referred to as mechanisms of delegated command. There are even provisions for commanders to act contrary to their written instructions when circumstances dictate.[27]

Although some scholars have recognized that these features exist in the United States military command and control system, the actual complexity of that system has not fully grasped in the literature on crisis management.[28] In the conduct of military operations, commanders at all levels in the chain of command have significant decisionmaking authority and can do much more than simply execute presidential policy decisions. An understanding of the mechanisms through which authority to make operational decisions is delegated to military commanders is essential

---

[27]See Chapter VII for a detailed description of military command and control.

[28]Two notable exceptions to this lack of awareness are John Steinbruner, "An Assessment of Nuclear Crises," in Franklin Griffiths and John C. Polanyi, eds., The Dangers of Nuclear War (Toronto: University of Toronto Press, 1980), pp. 34-49; Scott D. Sagan, "Nuclear Alerts and Crisis Management," International Security 9 (Spring 1985): 99-139.

for accurately assessing the crisis management problems that arise when military forces are employed in crises.

Returning to the organizational process model, its two serious flaws can now be seen. First, it fails to recognize that many government organizations, the armed forces in particular, are delegated siginificant authority to make detailed decisions on how to carry out policies.[29] The normal state of affairs--in crises as well as in peacetime and war--is for there to be innumerable military actions taking place that the President is not directly controlling. Second, the organizational process model fails to account for the fact that decisionmakers in many goverment organizations, particularly military commanders, often face circumstances that had not been anticipated by national leaders when deciding upon a course of action. Lower-level decisionmakers can be confronted with a require-ment to make what is essentially a policy decision without specific guidance on how to make it or sufficient time to seek further guidance from higher authority. Thus, the organizational process model must be modified to account

---

[29]Allison, pp. 85-6, recognizes that "Government action requires decentralization of responsibility and power," but mentions this only as being the reason why national leaders intervene in the internal processes of organizations dealing with military and foreign policy. His model recognizes only one legitimate policymaker, the President, and treats all other governmental actors as advisors or administrators without autonomous policymaking authority.

for the substantial legitimate decisionmaking authority routinely delegated to military commanders.

Replacement of the single interaction sequence model of crises in favor of a model recognizing the existence of multiple interaction sequences corrects many of these weaknesses in crisis theory. Each interaction sequence consists of a series of actions and reactions between specific groups of decisionmakers on each side. Although any number of interaction sequences could be postulated, limited only by the number of decisionmakers capable of affecting the crisis, this results in a model of excessive complexity. Instead, a relatively simple model of three interaction sequences will be used. Each of the three interaction sequences will be associated with a specific level in the chain of command, leading to a depiction of crisis interaction as being stratified into three levels. This will be referred to as the stratified interaction model. The stratified interaction model is described in detail in Chapter III.[30]

---

[30]Although this study is limited to interactions among military forces, the stratified interaction model can be applied to any organization that is sufficiently large and comples that top-level decisionmakers are incapable of exercising continuous direct control of its myriad interactions with the environment. Thus, in the U.S. Government, the stratified interaction model would apply to the Department of State and other large departments as well as to the Department of Defense. The model would also apply, for example, to large corporations and universities. The organization theory foundations of the stratified interaction model are discussed in Chapter IV.

The fundamental condition necessary for crisis interaction to be stratified is for the military establishment to be sufficiently large and complex that national leaders are incapable of exercising constant, direct, positive control of the actions of all operational units which might have an impact on the crisis. This condition is clearly met in the military establishments of the United States and the Soviet Union.[31] It can also be met in the military establishments of much smaller nations if national leaders do not have the capability or desire to exercise direct control of their forces.

When constant, direct, positive control of operational forces is not being exercised, different sets of decision-makers are delegated authority to make specified operational decisions. Their decisionmaking authority is bounded by the existing conditions of delegated command, which could range from being tightly controlled to being essentially autonomous. Even forces under the direct control of the

---

[31]Paul Bracken contends that this condition arises in the control of nuclear weapons: "In neither country [the United States and the Soviet Union] do leaders have the tight central control over nuclear arsenals offered in public relations statements. Instead, they rely on the vast organizations which are needed to manage the complex integration process. This has profound implications for maintaining political control over nuclear forces as they go on alert and operate in war." The Command and Control of Nuclear Forces (New Haven, CT: Yale University Press, 1983), p. 8. He also demonstrates that similar conditions pervade military command and intelligence systems--conventional as well as nuclear.

President retain a certain amount of decisionmaking authority, which can be substantial in some circumstances.

The fact that different sets of decisionmakers are responsible for making different operational decisions does not in itself lead to stratified interaction. Hypothetically, if all of those decisionmakers possessed identical beliefs, objectives, and perceptions, the operational decisions they make would be the same ones that national leaders would make if exercising positive control.[32]

Organization theory explains why this hypothetical situation will not necessarily be the case. Different organizations and sub-organizations possess distinct belief systems, referred to as an "organzational essences" or "bureaucratic ideologies" in organization theory, which shape the perceptions of their members. In military organizations, organizational belief systems become formalized in the strategic and tactical doctrines formulated for employment of their forces. Such doctrines typically vary widely among military organizations. Decisionmaking in organizations is bounded by cognitive limits on rationality, which generate a range of mechanisms for simplifying environmental complexity, coping with ambiguity, and dealing with value complexity. A principle

---

[32]This is the assumption that is made in the "unitary actor" model of national behavior. For example, see Bruce Bueno de Mesquita, The War Trap (New Haven, CT: Yale University Press, 1981), pp. 20-23.

effect of the cognitive limits on decisionmaking is to give

prominence to the beliefs and perceptions held by individual

decisionmakers. Thus, there are ample theoretical grounds

for expecting that military commanders will make operational

decisions different from those that national leaders would

have made if they had been in a position to make them.[33]

Paul Bracken's concept of "tightly coupled forces" is

an important contribution toward a more accurate understand-

ing of how the U.S. command and control system affects

crisis interaction. Bracken contends that U.S. and Soviet

nuclear forces are tightly coupled due to two features of

their respective command and control systems: vertical

integration of early warning sensors with operational

---

[33]On organizational ideology see Philip Selznick, TVA
and the Grass Roots (Berkeley: University of California
Press, 1953), pp. 47-50; Anthony Downs, Inside Bureaucracy
(Boston: Little, Brown and Co., 1967), pp. 237-46; Morton H.
Halperin, Bureaucratic Politics and Foreign Policy (Washing-
ton, DC: Brookings Institution, 1974), pp. 26-8. On
military doctrine as a belief system see Richard K. Betts,
Soldiers, Statesmen, and Cold War Crises (Cambridge, MA:
Harvard University Press, 1977), pp. 115-26.
On cognitive limits on decisionmaking see Herbert A.
Simon, "A Behavioral Model of Rational Choice," in Herbert
A. Simon, ed., Models of Man: Social and Rational (New York:
John Wiley and Sons, 1957), pp. 241-60; James G. March and
Herbert A. Simon, Organizations (New York: John Wiley and
Sons, 1958), Chapter 6; Herbert A. Simon, "Theories of
Decision-Making in Economics and Behavioral Science,"
American Economic Review 49 (June 1959): 253-83; Richard M.
Cyert and James G. March, A Behavioral Theory of the Firm
(Englewood Cliffs, NJ: Prentice-Hall, 1963), pp. 116-125;
John D. Steinbruner, The Cybernetic Theory of Decision: New
Dimensions of Political Analysis (Princeton, NJ: Princeton
University Press, 1974), pp. 88-139. On methods of coping
wirh cognitive limits, see George, Presidential Decision-
making, pp. 25-53.

nuclear forces on each side, necessary to reduce their

vulnerability to surprise attack, and a de facto coupling of

U.S. and Soviet forces through each side's warning and

intelligence networks.[34] Mutual coupling can drive an

interaction process between Soviet and American nuclear

forces:

> This mutual coupling occurs because a
> threatening Soviet military action or alert can be
> detected almost immediately by American warning and
> intelligence systems and conveyed to force
> commanders. The detected action may not have a clear
> meaning, but because of its possible consequences
> protective measures must be taken against it. The
> action-reaction process does not necessarily stop
> after two moves, however. It can proceed to many
> moves and can, and often does, extend from sea-based
> forces to air- and land-based forces because of the
> effect of tight coupling.[35]

This action-reaction process can produce what Bracken calls

"a mutually reinforcing alert," in which U.S. and Soviet

actions prompt increasingly higher alert levels on both

sides. A mutually reinforcing alert, in turn, would

exacerbate political tensions because of the near

impossibility of distinguishing precautionary military moves

from hostile political moves.[36]

Bracken's concept of tight coupling is an important

contribution to understanding crisis interaction, but it

does not convey the actual complexity of the relationships

---

[34]Bracken, pp. 54-65.

[35]Ibid., pp. 59-60.

[36]Ibid., pp. 64-65.

between American and soviet forces. Two modifications to
Bracken's concept are needed to derive the stratified
interaction model. First, national leaders are not
necessarily an integral element in all aspects of the
vertical integration of sensors with forces: some major
warning and intelligence systems are directly linked to the
commanders of operational military forces, who have been
been delegated authority to take certain actions on the
basis of warning provided by those systems without further
orders from national leaders. Bracken makes this clear in
his description of the command and control system, but
includes national leaders in the action-reaction loop when
describing the process of mutually reinforcing alerts. To a
degree, that process can proceed without national leaders
specifically having to order alerting actions as military
commanders act in compliance with their standing orders.

The second modification to Bracken's model is that
U.S. conventional forces are vertically integrated with
warning and intelligence systems, and tightly coupled with
Soviet conventional forces, in a manner similar to nuclear
forces. The reasons for this are the same: the ability of
U.S. conventional forces to successfully execute their
wartime missions can be crucially dependent on strategic
warning of an impending Soviet attack. This is most clear
in NATO, where the alliance defense strategy is based on
having sufficient warning to complete essential defensive

preparations. It is also true for the U.S. Navy, whose Maritime Strategy is founded on early and rapid surging of naval forces to key forward operating areas. An action by either side to increase the readiness of major conventional force commands is readily detected by the command on the other side responsible for dealing with that threat, which then takes actions to compensate for the changed strategic situation. This is the normal state of affairs in peacetime. As a crisis situation emerges the tightness of coupling between the conventional forces of the two sides actually increases as surveillance efforts are stepped up and focused on those adversary forces most likely to play an immediate role in the crisis.

In some military environments, particularly in naval warfare, U.S. and Soviet forces are tightly coupled down to the tactical level. At any given moment U.S. and Soviet tactical forces are operating in close proximity in several parts of the world: their naval forces routinely intermingle on the high seas, their ground forces are within sight of each other along the border between East and West Germany, and their surveillance aircraft are monitoring and being monitored by each other's air defense systems. Because a surprise attack by either side could be tactically decisive in an individual engagement, operational forces on both sides keep their adversary under close and constant surveillance. An action by either side's tactical forces to

increase their readiness or improve their tactical situation is readily detected by the other side's tactical forces in the vicinity, which then take actions to compensate for the changed tactical situation. Again, this is the normal state of affairs in peacetime, and the intensity of surveillance increases as the level of tensions rise in a crisis--further tightening the coupling between the forces of the two sides.

Thus, the actual situation is that the tight coupling of U.S. and Soviet forces is stratified into tight coupling at two levels: the major command level, including strategic nuclear forces and other major commands, and the tactical level, encompassing operational units in close proximity in the field or at sea. Significant information on the status of the other side's forces flows directly to military commanders at these levels from organic sensors under their control and dual reporting from intelligence sources outside their commands. Military commanders are only partially dependent on the chain of command to tell them what the adversary is doing, and are delegated authority to take certain specified measures to adapt the readiness of their forces to changes in the adversary's forces. Military commanders are obligated to immediately report such actions to their superiors, thus allowing their orders to be countermanded, if necessary.[37] The key point is that within

---

[37]See Chapter IV for a detailed discussion of military command and control.

specified limits, control of U.S. operational forces is delegated widely to commanders with their own sources of intelligence on Soviet forces.

In summary, interaction between the two sides in crises has been viewed in terms of an implicit single interaction sequence model that does not accurately describe the complexity of crisis interaction. The weakness of the single interaction sequence model is that in subtle ways it leads toward an implicit assumption that national leaders are in control of the actions taken by their nation in a crisis. The bureaucratic politics and organizational process models do not entirely correct this weakness and have serious problems of their own. The key to correcting these weaknesses is a more accurate understanding of the complexity of the military command and control system, in which military commanders are delegated significant decisionmaking authority.

A model containing three interaction sequences--each sequence associated with a specific level in the chain of command--will be used. Crisis interaction is stratified when the military establishment of a country is sufficiently large and complex that national leaders are incapable of exercising direct control over all operational units that could have an impact on a crisis. Paul Bracken's concept of "tightly coupled forces"--modified by the observation that tight coupling is stratified, occurring separately at the

strategic and tactical levels--then explains how interaction sequences can arise.

Crisis Management

A nation confronted by a crisis can choose from among three general strategies for dealing with it: capitulation, war, or crisis management. As defined by Williams, "crisis management is concerned on the one hand with the procedures for controlling and regulating a crisis so that it does not get out of hand and lead to war, and on the other hand with ensuring that the crisis is resolved on a satisfactory basis in which the vital interests of the state are secured and protected."[38] These two elements are also central to the definition of crisis management used by Snyder and Diesing: first, exercise of detailed control by the top leadership in order to to avoid war, and, second, efforts by national leaders "to advance or protect their state's interests, to win or at least to maximize gains or minimize losses, and if possible to settle the issue in conflict so that it does not produce further crises."[39] This is the definition of crisis management that will be used in this study.

The essence of the crisis management problem is to find the optimum balance between efforts to advance or

---

[38]Williams, p. 30.

[39]Snyder and Diesing, p. 207.

protect national interests, and efforts to avoid war.
Williams describes crisis management as an attempt to
balance attainment of national goals in the bilateral
competition against efforts to avoid the shared danger of
war.[40] Similarly, Snyder describes crisis management as
balancing coercion against disaster avoidance, and balancing
accommodation against loss avoidance.[41] This conceptualiza-
tion of crisis management as balancing between pursuit of
national interests and avoidance of war will be a foundation
for the concept of political-military tensions to be
developed in Chapter III.

The ability of national leaders to maintain control
over events is a central problem in crises. Decisionmakers
commonly perceive that a crisis can develop a self-
sustaining force or impetus of its own, degrading their
ability to control events.[42] According to Thomas C.
Schelling, "It is the essence of a crisis that the
participants are not fully in control of events; they take
steps and make decisions that raise or lower the danger, but
in a realm or risk and uncertainty."[43] Thus, maintaining

---

[40]Williams, p. 29.

[41]Glenn H. Snyder, "Crisis Bargaining," p. 240. Also
see Snyder and Diesing, p. 270.

[42]Young, pp. 19-20; Williams, p. 26.

[43]Thomas C. Schelling, Arms and Influence (New Haven,
CT: Yale University Press, 1966), p. 97.

control over events also means maintaining control of risks, particularly the risk of war breaking out inadvertently.

On the other hand, Schelling's description of brinkmanship as "manipulating the shared risk of war," and his concept of "the threat that leaves something to chance" do not convey an accurate image of how national leaders manage crises.[44] Snyder has observed that, while delegating control of military operations to subordinate commanders (or threatening to do so) can be used as a coercive tactic in crisis bargaining, national leaders normally emphasize maintaining direct control of military forces in order to avoid war.[45] Thus, Schelling's concepts of manipulation of risk and the threat that leaves something to chance are better viewed as interpretations of the nature of crisis interaction, rather than as strategies consciously employed by national leaders in crises.

National leaders can be confronted with serious problems in attempting to maintain control over events in a crisis. Glenn Snyder identifies four "autonomous risks" that could cause a loss of control over events: military action being driven by its own logic and momentum, national leaders losing control over their military commanders, lack of military options other than escalatory war plans, and

---

[44]Ibid., pp. 99-105; Schelling, Strategy of Conflict, pp. 187-203.

[45]Glenn H. Snyder, "Crisis Bargaining," pp. 244-245.

impairment of rational calculation by psychological factors under the stress of a crisis.[46] This study will treat these problems somewhat differently, but Snyder's list of autonomous risks is a useful summary of the types of concerns that have been raised in the crisis managment literature.

A wide range of actions can be taken with military forces during a crisis in pursuit of military and political objectives. Coral Bell identifies "signals," threats or offers communicated to the other side, as the basic instrument of crisis management, and notes that some of the most effective signals are movements of military resources.[47] As Bell suggests, political signalling is a primary function of military forces in crises, competing with or even overshadowing their nominal military missions. Alexander L. George lists five general uses to which military forces can be put in crises: reducing the vulnerability and increasing the readiness of theater and strategic nuclear forces, signaling limited intentions and an interest in avoiding escalation, engaging in a test of military capabilities within restrictive ground rules, conveying military threats for coercive pressure in bargaining, and deterring escalation by the adversary and

---

[46]Ibid., p. 241.

[47]Bell, p. 73.

neutralizing his coercive threats.[48] These categories show the ways in which military forces serve both political and military purposes in crises.

Although attempts have been made to draw distinctions between actions taken for military purposes and actions taken for political purposes, virtually all military actions undertaken in a crisis have a dual political-military nature. Actions taken for military purposes, such as increasing the readiness or reducing the vulnerability of military forces, can have political impact if perceived as a signal of hostile intent. Conversely, actions taken for political purposes, such as withdrawing forces from a contested area to signal limited objectives or increasing forces in a contested area to apply coercive or deterrent pressure, can have military impact by shifting the local balance of forces and altering the capabilities available to local military commanders. In a crisis, political missions such as coercion and signalling intentions are assigned to forces that must also be ready for limited combat operations and the possibility of sudden escalation to full-scale war.

Studies of crisis management have identified stringent requirements for its success. Foremost among these, as Bell points out, are imposing limits on the military means

---

[48]Alexander L. George, "Crisis Management: The Interaction of Political and Military Considerations," Survival 26 (September/October 1984): 229-33.

employed that are commensurate with the limited ends of
military action in crises, and maintaining close diplomatic
control of military measures.[49] The basic requirement that
national leaders maintain close control of military
operations is central to the more detailed lists of
requirements and techniques proposed in other studies.

In addition to the "political" requirements of crisis
management--limiting objectives and the means employed to
secure those objectives--George also identifies seven
"operational" requirements for crisis management: First,
political authorities must control military operations,
including details of deployments and low-level actions as
well as selection and timing of the moves. Second, the
tempo of military operations may have to be deliberately
slowed, creating pauses for the exchange of diplomatic
signals, assessment, and decisionmaking. Third, military
actions have to be coordinated with diplomatic actions in an
integrated strategy for resolving the crisis acceptably
without war. Fourth, military actions taken for signalling
purposes must send clear and appropriate signals consistent
with diplomatic objectives. Fifth, military options should
be avoided that give the adversary the impression of an
impending resort to large-scale warfare, possibly prompting
him to pre-empt. Sixth, military and diplomatic options

---

[49]Bell, 49.

should be chosen that signal a desire to negotiate a
solution to the crisis rather than to seek a military
solution. Seventh, military options and diplomatic
proposals should leave the adversary a way out of the crisis
compatible with his fundamental interests.[50] These crisis
management requirements have important implications for
manner in which military force is used and controlled in
crises.

Ole R. Holsti has identified six crisis management
techniques, four of which address the use of force in
crises: First, avoiding steps that seal off "escape
routes." This precludes military actions which the other
side would perceive as leaving it no way out of the crisis
other than war. This technique places limits on military
options, calling for carefully limited use of force as
opposed to drastic, precipitous military actions. Second,
orchestrating actions, particularly military actions, with
declarations of intent so as to use multiple channels of
communication to convey the same message. This also affects
the employment of military forces, requiring that their
actions be coordinated with diplomatic moves for signaling
purposes. Third, making efforts to slow the pace of crisis

---

[50]George, "Crisis Management," p. 226. An earlier,
slightly different version of this list appeared in h's "The
Development of Doctrine and Strategy," in Alexander L.
George, David K. Hall, and William E. Simons, The Limits of
Coercive Diplomacy (New York: Little, Brown and Co., 1971),
pp. 8-11.

events. This requires dampening the tempo of military operations, which typically emphasize speed of execution for tactical success. Fourth, keeping responsible policy makers in control of the details of implementation as well as broad strategic decisions. This raises the civil-military relations issue of who is to control execution of military operations in the field, and whether political or military considerations should govern operational decisions.[51]

The requirements and techniques identified by George and Holsti are similar. Both emphasize close control of military operations by national leaders, and tailoring of military options to support crisis bargaining and avoid escalation of the crisis.

A significant weakness in the crisis managment literature is that, with few notable exceptions, there has been scant recognition that tensions can arise in attempting to reconcile military considerations with crisis management requirements. This arises from the nature of crisis management: the objective is to protect vital national interests as well as to avoid war, and military force is being employed for signaling and coercion. Secretary of Defense Robert S. McNamara contributed to, and may have originated, the lack of attention to the military dimension of crisis management when he asserted in the wake of the

---

[51]Holsti, pp. 221-226.

Cuban Missile Crisis that "Today there is no longer any such thing as strategy, there is only crisis management."[52] McNamara overlooked that there are requirements for effective employment of military force just as there are requirements for effective crisis management, and conflicts may arise between the two sets of requirements.

Alexander L. George rejects McNamara's antithesis between strategy and crisis management, observing that "in reality, policy-makers need to employ broad strategic principles to help them to reconcile and integrate, however imperfectly, the often competing requirements of force and diplomacy."[53] Although he makes it clear that political considerations are paramount and that close presidential control of military operations is crucial for effective crisis management, George points out that "there are likely to be severe limits on the ability of top-level political authorities to orchestrate military operations and serious risks if they attempt to carry 'micro-management' of military forces too far."[54] This is an crucial point that has received little attention in crisis management studies.

The nature of the tensions that can arise between political and military considerations in a crisis have been

---

[52]Quoted in Bell, p. 2.

[53]George, "Crisis Management," p. 224.

[54]Ibid., p. 233.

described by Eliot A. Cohen in an assessment of the Cuban
Missile Crisis:

> The events of October 1962 created considerable
> tension between military men seeking to protect those
> under their command, in the event of an outbreak of
> war, and politicians seeking to give the other side
> time to think and give in. Had men in fact died as a
> result, had ships sunk or airplanes fallen by the
> score, the crisis in civil-military relations would
> have taken a more dramatic turn, one in which, I
> suspect, civilian leaders would have accommodated
> commanders far more than they actually did.[55]

Thus, the weakness in the crisis management literature is
that it has not recognized that important, legitimate
military considerations arise when military forces are
employed as a political instrument in crises, and that
tensions can arise in attempting to reconcile military
requirements with crisis management requirements. Crisis
management did not replace military strategy--other than in
the minds of some social scientists--it created complex
challenges for effective formulation and execution of
military strategy on behalf of political-diplomatic
objectives.

In summary, crisis management is the exercise of
detailed control of diplomatic and military activities by
national leaders in order to to avoid war while attempting
to advance their state's interests or protect those
interests against losses during a crisis. The essence of
the crisis management is to find the optimum balance between

---

[55]Eliot A. Cohen, p. 6.

efforts to advance or protect national interests, and efforts to avoid war. Political signalling is a primary function of military forces in crises. Virtually all military actions undertaken in a crisis have a dual political-military impact: sending political signals and affecting the balance of military capabilities. Foremost among the requirements for the success of crisis management are imposing limits on the military means employed which are commensurate with the limited ends being sought, maintaining close top-level control of military measures, and carefully tailoring military options to support crisis bargaining and avoid escalation. A serious weakness of the crisis management literature is that it has not adequately addressed the tensions that can arise between these crisis management requirements and military considerations.

Crisis Stability

The third weakness in the crisis management literature is that the concept of crisis stability is poorly developed and there is a poor understanding of the escalation processes that could cause a crisis to escalate to war. Crisis stability is viewed as being primarily a function of weapons technology, particularly the degree to which it gives an advantage to the offense, and military doctrine, particularly doctrines emphasizing the superiority of the offensive. Lacking is an appreciation of the operational

factors that affect crisis stability once a decision is made to employ military forces in a crisis. The escalation processes that could cause a crisis to escalate to war are also poorly developed. There is growing concern over inadvertent or accidental war, but these concepts are not well defined and there is a very low probability that any of the scenarios would occur.

The definition of crisis stability generally accepted in the crisis management literature is that crisis stability exists when neither side has an incentive to strike the first military blow, launching a preemptive attack on the other side. Alexander George adds a second dimension: crisis stability exists when neither side perceives that crisis management had broken down and cannot be restored. Thus, crisis stability is a function of the strategies each side is pursuing in the crisis as well as a function of weapons technology. This will be discussed further below.

The concept of crisis stability has generally been used to assess the stability implications of particular weapons technologies and force postures. Weapons that enhance crisis stability are survivable, providing an assured retaliatory capability, and do not provide first strike capabilities for use against the other side. Weapons that degrade crisis stability are vulnerable to preemption, potentially confronting leaders with a "use them or lose them" dilemma in a crisis. The most destabilizing weapons

are those which are valuable for launching a first strike--
such as by providing a rapid, precise hard-target kill
capability--but which are themselves vulnerable to
preemption. Crisis stability as a technological
characteristic is also applied to command and control
systems: survivable systems enhance stability by ensuring
that retaliation can be executed, while vulnerable systems
degrade stability by providing the other side an incentive
to preempt for damage limitation purposes.[56]

Although this definition of crisis stability is useful
for assessing weapons and force postures, it is too narrowly
focused on technology for the purposes of this study. What
is needed is a broader definition encompassing the full
range of factors which could cause efforts at crisis manage-
ment and escalation control to fail, resulting in war.

The concept of the security dilemma, originally
proposed by Herbert Butterfield, provides a useful
foundation for defining a broader concept of crisis

---

[56]Bernard Brodie, Strategy in the Missile Age
(Princeton, NJ: Princetion University Press, 1959), pp. 300-
303; Lebow, Between Peace and War, pp. 238-242; Holsti, pp.
228-233; and Schelling, Arms and Influence, pp. 224-225, 234-
235. On the role of command and control system vulner-
ability in the security dilemma, see Phil Williams, "Crisis
Management: The Role of Command, Control and Communica-
tions," RUSI Journal 128 (December 1983): 33-39; Garry D.
Brewer and Paul Bracken, "Some Missing Pieces of the C^3I
Puzzle," Journal of Conflict Resolution 28 (September 1984):
451-469; and Bruce G. Blair, Strategic Command and Control:
Redefining the Nuclear Threat (Washington, DC: Brookings
Institution, 1985), pp. 284-285.

stability.[57] The security dilemma, as defined by Robert Jervis, is that "many of the means by which a state tries to increase its security decrease the security of others."[58] According to Jervis the intensity of the security dilemma is a function of three factors: (a) the condition of anarchy in international politics, in which states tend to pursue security unilaterally rather than accept the risks of cooperation with potential adversaries; (b) geography, commitments, and beliefs, which can create the perception that the security of the state and its interests (such as territories abroad, commerce, and allies), requires the ability to take offensive action against others; and (c) the perception that military technology and geography give offense a strategic advantage over defense, which can be exacerbated by difficulty in distinguishing defensive from offensive weapons. When a decisive advantage can be gained by striking first, such as when military forces are vulnerable to preemption, even a status quo power without expansionist objectives has an incentive to strike first.[59]

The security dilemma is used by Butterfield and Jervis primarily to explain how arms races and international

---

[57]Herbert Butterfield, History and Human Relations (London: Collins, 1951), pp. 19-20.

[58]Robert Jervis, "Cooperation Under the Security Dilemma," World Politics 30 (January 1978), p. 169.

[59]Ibid., pp. 167-214.

tensions arise from unilateral efforts by states to protect their security and maintain the balance of power. But Jervis suggests it also applies to crisis stability: "The second aspect [of the offense-defense balance]--whether it is better to attack or defend--influences short-run stability. When the offense has the advantage, a state's reaction to international tension will increase the chances of war." The reason for this is that when there are incentives for preemption and reciprocal fear of surprise attack, "There is no way for the state to increase its security without menacing, or even attacking, the other."[60] In Jervis' view, this problem arises from the existence of the security dilemma as a feature of international politics, as opposed to being a phenomenon unique to crises.

Before applying the security dilemma to crisis stability, a expansion of Jervis' definition is needed. Reciprocal fear of surprise attack and incentives for preemption arise from three sources: the perceived impact of weapons technology on the nature of warfare, perceptions of the adversary's military strategy and doctrine, and the operations being conducted by military forces.[61]

---

[60]Ibid., p. 188.

[61]This is derived from Schelling's description of how weaponry influences the nature of crises and the processes by which wars start: "To impute this influence to 'weaponry' is to focus too narrowly on technology. It is

The perceived impact of weapons technology on the nature of warfare exacerbates the security dilemma in crises. The offense-defense balance applies to all aspects of warfare: conventional ground, air, and naval warfare, as well as strategic nuclear warfare. When the prevailing weapons technologies in a particular area of warfare are perceived as giving an inordinate advantage to offensive action or being the first to strike, military commanders will have a strong incentive to preempt. The offense-defense balance varies across warfare areas: as will be explained in the next chapter, naval warfare is especially offense-dominant, resulting in great stress being placed on striking first.

The perceived impact of weapons technology on the offense-defense balance is not the only factor exacerbating the security dilemma, perceptions of military strategy and doctrine are equally important. In fact, the difficulty of distinguishing offensive from defensive weapons tends to make strategy and doctrine _more_ important than technology. Most weapons, including virtually all conventional weapons not emplaced in fixed fortifications, can be used with nearly equal effectiveness for offense or defense. Their offensive or defensive nature is predominantly a function of

---

weapons, organization, plans, geography, communications, warning systems, intelligence, and even beliefs and doctrines about the conduct of war that together have this influence." _Arms and Influence_, p. 234.

the military strategy and doctrine prescribing how those
weapons will be used in war.

In assessing the threat posed by a potential enemy's
forces and in making contingency plans against that threat,
decisionmakers attempt to estimate the adversary's
intentions.[62] Estimating intentions, in turn, requires
either estimating or make assumptions about how the
adversary would use its forces in wartime, which is the
essence of military strategy and doctrine. Thus, the
security dilemma can arise from perceptions held by each
side that the other side has adopted an offensive military
strategy or a military doctrine emphasizing preemption or
surprise attack. In circumstances of mutual perceptions of
offensive strategies, actions taken by each side to increase
its security, even when motivated by defensive intentions,
will be perceived by the other side as decreasing its
security.[63] This idea in implicit in Jervis' definition of

---

[62]J. David Singer, "Threat Perception and the Armament
Tension Dilemma," Journal of Conflict Resolution 2 (March
1958): 93-94; Dean G. Pruitt, "Definition of the Situation
as a Determinant of International Action," in Herbert C.
Kelman, ed., International Behavior: A Social-Psychological
Analysis (New York: Holt, Rinehart and Winston, 1965),
p.400; Raymond Cohen, Threat Perception in International
Crisis (Madison: University of Wisconsin Press, 1979), p.5.

[63]The worst case from a crisis stability perspective
is when both sides in a conflict have adopted offensive
strategies, and accurately perceive that the other side has
adopted an offensive strategy. Under these circumstances
each side has an incentive to strike first so as to be able
to effectively execute its own strategy and preempt the
enemy from executing his. The "cult of the offensive"

the security dilemma, but is subordinated to his emphasis on technology.

The third source of the security dilemma is the operations being conducted by military forces.[64] In his study of threat perception, Dean G. Pruitt identifies military actions as an important source of evidence used to infer the intentions of an adversary.[65] That military operations are used as an indicator of intent is, of course, the basis for using military forces for signalling in crises. But crisis military operations also help to define whether the forces being employed have an offensive or defensive purpose. Military operations can thus perform the same function as strategy and doctrine: defining or signalling the offensive or defensive nature of forces that whose technological characteristics make them suitable for either role. Naval vessels, for example, may appear defensive when kept close to their homeports, far from the scene of a crisis, but appear offensive to an adversary when deployed off his coast. This can occur regardless of the

---

school of thought contends that this was the strategic environment in 1914, when a relatively minor incident rapidly escalated to war. See Van Evera, pp. 58-107; and Jack Snyder, "Civil-Military Relations," pp. 108-146

[64]Scott Sagan refers to this as "the operational dimension of crisis stability." See Sagan, "Managing Strategic Nuclear Alert Operations," (Santa Monica, CA: Rand, forthcoming, 1988).

[65]Pruitt, pp. 403-404.

intention of the deployment, which could well be defensive or deterrent. Similarly, forward deployed naval vessels can appear to have offensive purposes even when the national strategy they support is essentially defensive or deterrent. Military operations can also reinforce perceptions of strategy and doctrine, appearing to confirm estimates or assumptions that an adversary holds an offensive strategy.

To apply the security dilemma idea to analysis of crisis stability, this study will use the concept of the crisis security dilemma: In a crisis, many of the actions a state takes to increase its security and improve its bargaining position decrease the security of the adversary, at least in his perception. This is a particular case of the security dilemma as defined by Jervis. The primary difference is that in a crisis the most important sources of the dilemma are the military strategies and doctrines of the two sides, and especially the military operations being conducted by the two sides. Under normal (non-crisis) peacetime conditions, the "many actions" Jervis refers to are primarily force posture and weapons procurement decisions, in which the technological characteristics of the forces play an important role in determining the offense-defense balance. In a crisis, decisionmakers focus on the adversary's immediate intentions and the actions he is taking with his military forces, making these factors predominant in determining the offense-defense balance. The

implication is that the security dilemma can be much more severe in a crisis, when military forces are being used for coercive and deterrent threats.

A second aspect of crisis stability that is not well developed in the existing literature is escalation from crisis to war. Concepts that have been used to address this topic include the escalation spiral and accidental or inadvertent paths to war. Although some of these concepts are useful, they have not been well integrated with other crisis and crisis management concepts. Even the point at which a confrontation shifts from being a crisis to being a war is unclear in the literature. We lack a separate term to describe the transitional state of conflict that exists during the period after violence erupts but before a limited war exists. Some analyses implicitly limit crises to political disputes in which use of military force is only threatened, not actually carried out. Other analyses encompass the use of force, such as to achieve a military fait accompli in a crisis. Both approaches have merit, but for the purposes of this study the definition of crisis will include limited use of force as well as the threat of force.

The distinction between crisis and limited war will be based on the perceptions and strategies held by national leaders on the two sides. If they perceive themselves as involved in a crisis or as attempting to prevent a conflict from erupting in war, then the conflict is a crisis even if

fighting has broken out. If they perceive themselves as launching or fighting a limited war, then the conflict has transitioned to a state of war, no matter how limited.

There are potential problems with this approach. The point at which a confrontation shifts from being a crisis to being a war could be difficult to ascertain in actual cases, and even be unclear in the minds of leaders on the two sides. Nations can be involved in a "phony war," in which there is a declared state of war but no fighting, as were Britain and France with Germany from September 1939 to May 1940. Nations can also be involved in recurring episodes of intense but brief fighting without there being a declared state of war, as were the Soviet Union and Japan along the Manchurian border from July 1938 to September 1939, and the Soviet Union and the People's Republic of China in 1969. Nevertheless, basing the distinction between crisis and war on the perceptions of the participants is superior to an arbitrary definition of crisis that excludes the use of force.

Several studies of conflict and war have proposed that an escalation spiral can cause tensions and insecurities to erupt in war.[66] In a refinement of this theory, Richard Smoke concludes that there is an escalation dynamic driven

---

[66]For a discussion of spiral theories of escalation, see Robert Jervis, <u>Perception and Misperception in International Politics</u> (Princeton, NJ: Princeton University Press, 1976), pp. 58-113.

by rising stakes in the outcome of a conflict and an action-reaction cycle. Rising stakes increase the motivation of national leaders to prevail in the crisis. In the action-reaction process an escalatory action by one side provokes an escalatory reaction by the other side in recurring cycles.[67] Although Smoke's analysis is limited to the escalation processes that occur after war has broken out, it is equally applicable to the escalation processes that can arise after fighting erupts in a crisis.

The escalation spiral that led to the outbreak of World War I is often cited as the classic example of escalation dynamics at work. Bell has described the 1914 case as being an example of a "crisis slide," in which a series of crises gather irresistible momentum toward war. The escalatory impact of a crisis slide is that "the decisionmakers of one or more of the dominant powers believe that they see the options available to them steadily closing down to the single option of war or unlimited defeat."[68] The events of 1914 have also led to the view that inflexible war plans and offensive military doctrines can create a strategic environment in which national leaders are unable to control the momentum of events and seek a diplomatic

---

[67]Richard Smoke, War: Controlling Escalation (Cambridge, MA: Harvard University Press, 1977), pp. 23-35, 268-297. Also see Williams, pp. 97, 101.

[68]Bell, pp. 14-15.

solution to a crisis.[69] Thus, the danger of escalation is not limited to the effects of individual events, but includes the danger of an uncontrollable escalatory cycle leading to wa:.

Understanding the events or pressures that can trigger an escalatory spiral is at least as important as understanding the dynamics that drive the spiral after it starts. Glenn Snyder has identified four "autonomous risks" that could trigger uncontrollable escalation: military action being driven by its own logic and momentum, national leaders losing control over their military commanders, lack of military options other than escalatory war plans combined with pressure to take action, and psychological factors impairing rational calculation under the stress of a crisis. Of these, Snyder views psychological factors impairing rational calculation as most likely and losing control over military commanders as least likely.[70] All of these factors are compatible with Smoke's theory of escalation dynamics.

Several possible paths to war have been proposed. The basic categories are premeditated attack, catalytic war, accidental war, preemption, and inadvertent war. A premeditated war is launched deliberately, usually (but not

---

[69]Van Evera, pp. 63-65, 71-79; Jack Snyder, "Civil-Military Relations," pp. 112-114, 125-129.

[70]Glenn H. Snyder, "Crisis Bargaining," p. 241.

always) by surprise attack, and is often described as a "bolt from the blue" attack. Catalytic war is one started by a third party, which can be either a nation or some other group, such as a terrorist organization. The typical scenario is the launching of a nuclear weapon at one of the two superpowers, which responds by retaliating against the other superpower thinking it to be the source of the initial blow. Accidental war is the result of either equipment malfunctions or unauthorized use of nuclear weapons by military commanders. Equipment malfunctions can occur in strategic warning systems, providing false warning of an attack; in command and control or battle management systems, again providing false indications of attack or spurious orders to launch an attack; or in nuclear weapons and their control systems, resulting in accidental launch of nuclear weapons.[71] These three paths to war are generally regarded as much less likely than the other two. Additionally, although these three paths to was can occur whether or not a crisis is in progress, they are probably more likely to occur in crises as military forces are alerted for readiness and political signalling purposes.

---

[71]Thomas C. Schelling and Morton H. Halperin, Strategy and Arms Control (New York: Twentieth Century Fund, 1961), pp. 10-17; Herman Kahn, On Escalation (New York: Praeger, 1965), pp. 284-6; Graham T. Allison, Albert Carnesale, and Joseph S. Nye, Jr., "Introduction," in Graham T. Allison, Albert Carnesale, and Joseph S. Nye, Jr., eds., Hawks, Doves & Owls: An Agenda for Avoiding Nuclear War (New York: W.W. Norton, 1985), pp. 10-13.

The remaining two paths to war are particularly relevant to the study of crises. Preemption is motivated by perceptions and fears that the other side is about to strike first.[72] This is the path to war that results from the crisis security dilemma. Preemption can, of course, also result from correct perceptions that the other side is about to launch a premeditated attack. There are thus two preemption paths, one generated by the crisis security dilemma and the other generated by an actual impending premeditated attack. They are much different in terms of the analytical questions they raise: the first focuses on the dynamics of the security dilemma, while the second focuses on deliberate decisions to resort to war rather than continue crisis management. This study will address the preemption path that arises from the crisis security dilemma.[73]

Inadvertent war arises from an escalation process in which the two sides employ increasingly threatening military and diplomatic moves--including alerts, mobilizations, deployments of forces, small-scale demonstrative use of conventional weapons, and ultimatums--in an effort at gaining leverage in crisis bargaining and improving their

---

[72]Allison, Carnesale and Nye, pp. 10-13.

[73]Derived from Paul Bracken, "Accidental War," in Graham T. Allison, Albert Carnesale, and Joseph S. Nye, Jr., eds., Hawks, Doves & Owls: An Agenda for Avoiding Nuclear War (New York: W.W. Norton, 1985), pp. 29-37.

military positions in the event diplomacy fails. Accidents and other inadvertent military actions can contribute to this process. Such deliberate and inadvertent actions increase tensions and harden resolve (similar to the manner described by Smoke) until the process results in a war that neither side wanted or expected when the crisis first arose. This is a useful concept, but suffers from insufficient specificity as to how the individual actions contribute to an escalation process and omits significant factors that can also contribute to inadvertent war.

There are two weaknesses in the inadvertent war concept. First, it does not directly address the nature of the decision for war that arises out of the escalation process. There is just an "unintended eruption" of war, in Bracken's words.[74] Even in the inadvertent war scenario, the decision for war falls into one of two categories: deliberate or preemptive. A deliberate decision for war could result from the perception that the other side cannot be bargained or coerced into making the concessions being demanded, leaving war as the only perceived means for avoiding severe damage to vital national objectives. The distinction between this type of inadvertent war and premeditated war is that in the inadvertent path the deliberate decision for war is made under the stress of a

---

[74]Ibid., p. 29.

crisis, after an escalatory process defeats crisis bargaining. A decision for preemptive war could result from the crisis security dilemma--the escalatory process generates perceptions that the adversary is preparing to strike first.

The second weakness in the inadvertent war concept is that it does not capture the true complexity of the crisis escalation process. In a crisis, interaction is stratified into multiple interaction sequences that can evolve semi-independently of each other. It is theoretically possible for fighting to erupt and an escalation process to be set in motion between the forces of the two sides at the scene of a crisis without escalation occurring in other interaction sequences between the two sides. There appear to be factors that inhibit the crisis escalation process from occurring and inhibit inadvertent war from resulting even when escalation does occur in a crisis. Identifying those factors will make a significant contribution to our understanding of the inadvertent war path and the strengths and weaknesses of crisis management.

In summary, the crisis security dilemma is that in a crisis, many of the actions a state takes to increase its security and improve its bargaining position decrease, or can appear to decrease, the security of the adversary. The most important sources of the dilemma are the military doctrines and the military operations being conducted by the

two sides. The definition of crisis used in this study will
include limited use of force as well as the threat of force,
with the distinction between crisis and limited war based on
the perceptions held by national leaders on the two sides.
There is an escalation dynamic driven by rising stakes in
the outcome of a conflict, which increase the motivation of
national leaders to prevail, and an action-reaction process,
in which an escalatory action by one side provokes an
escalatory reaction by the other side in recurring cycles.

The preemption and inadvertent paths to war are
particularly relevant to the study of crises. Preemption is
motivated by perceptions that the other side is about to
strike first. The preemption path to war that results from
the crisis security dilemma is the path that will be
addressed in this study. Inadvertent war arises from an
escalation process in which the two sides employ
increasingly threatening military and diplomatic moves in an
effort at gaining leverage in crisis bargaining and
improving their military positions. Inadvertent military
incidents contribute to this process. The escalation
process increases tensions and hardens resolve until it
results in a deliberate or preemptive decision for war.

Misperception in Crises

One of the most difficult problems of crisis
management is avoiding misperceptions of intentions and

objectives. Misperceptions can affect crisis management in three ways. First, they can erode the credibility of deterrent threats. Second, they can defeat attempts to signal limited objectives and a desire to resolve the conflict without war. Third, they can exacerbate the problem of the crisis security dilemma.

Deterrent threats often play a major role in crisis management. The effectiveness of a deterrent threat is dependent upon its credibility. For a variety of military, political, and cognitive reasons, the nation to be deterred may not perceive the deterrent threat as being credible, or may miscalculate the consequences of challenging a deterrent threat, leading to a failure of deterrence.[75] The credibility of extended deterrence can be particularly difficult. Once again, a variety of factors can cause an adversary to doubt the credibility of a commitment to defend an ally or client.[76] The relevant points for this study are that national leaders are generally concerned about the

---

[75]William W. Kaufmann, The Requirements of Deterrence (Princeton, NJ: Center for International Studies, 1954), pp. 6-8; Henry A. Kissinger, The Necessity for Choice (new York: Harper, 1960), pp. 40-41; Alexander L. George and Richard Smoke, Deterrence in American Foreign Policy: Theory and Practice (New York: Columbia University Press, 1974), p. 64; Schelling, Arms and Influence, pp. 36-43.

[76]Schelling, Arms and Influence, pp. 49-50; George and Smoke, pp. 550-65; Lebow, Between Peace and War, pp. 82-90. On extended deterrence see Brodie, pp. 252-55; and Warner R. Schilling, et al., American Arms and a Changing Europe (New York: Columbia University Press, 1973), pp. 5-15.

credibility of commitments and deterrent threats, and that in spite of their efforts to enhance the credibility of deterrent threats, failures of deterrence can occur for reasons beyond their control.

Concerns over credibility and misperception affect the use of military force as a political instrument in crises: the role of military forces is often to enhance the credibility of deterrent threats, but the threats those forces are intended to convey may be misperceived or otherwise fail to deter. Due to credibility problems with extended deterrent threats, particularly when threatening punishment by nuclear retaliation, many actions taken with military forces in crises are intended to enhance the credibility of extended deterrence by adding a specific threat of denial with conventional forces to the standing threat of punishment with strategic nuclear forces.[77]

Misperception of the intentions of an adversary, and miscalculation of the costs he is willing to endure or capable of exacting, can arise from several sources: the normal cognitive constraints on decisionmaking, from the particular psychological factors that affect decisionmaking

---

[77]On the concept of deterrence by denial and punishment, see Glenn H. Snyder, <u>Deterrence and Defense</u> (Princeton, NJ: Princeton University Press, 1961), pp.14-16, and <u>Deterrence by Denial and Punishment</u> (Princeton, NJ: Center for International Studies, 1959), p. 1; George H. Quester, <u>Deterrence Before Hiroshima</u> (New York: John Wiley and Sons, 1966), p. 2-4.

under stress, from the political and organizational
perspectives of participants in decisionmaking, and from
incomplete or inaccurate information on the adversary and
the status of the conflict.[78] Although national leaders
often make efforts to anticipate how adversary decision-
makers will perceive various crisis moves, attempting to
predict perceptions and reactions is inherently the weakest
aspect of crisis management. Thus, careful attention to the
clarity of signals being sent to the adversary may not
suffice to prevent escalation of the conflict.

The sources of misperceptions and their general role
in crises are well developed in the crisis management
literature. However, the effect that decisionmaker
awareness of the danger of misperception has on decisions
concerning the use of force in crises remains a weak point.
The danger of signals sent by military forces being
misperceived creates a dilemma for decisionmakers attempting
to use force as a political instrument in a crisis. The
misperception dilemma, as this problem will be called,
pervades all decisions on the use of force in a crisis.
There are actually two misperception dilemmas: the first
affects signals to adversaries, and the second affects
signals to allies and friends. The dilemma in signaling

---

[78]Charles Lockhart, Bargaining in International
Conflicts (New York: Columbia University Press, 1979), pp.
37-87; Holsti, pp. 7-25; Snyder and Diesing, pp. 340-418;
Lebow, Between Peace and War, pp. 148-228.

adversaries is between inadvertent signals of acquiescence and inadvertent signals of hostility. This misperception dilemma is present in both of the categories of crises: a direct crisis between the United States and another nation, and an indirect crisis arising from a conflict between two nations, one of whom is an ally or friend of the United States. The dilemma in signaling allies and friends is between inadvertent signals of encouragement and inadvertent signals of retrenchment. This misperception dilemma only arises in indirect crises, in which the U.S. role arises from its support for an ally or friend.

Efforts to signal limited objectives and interest in a negotiated solution, and to limit the level of tension and violence in a crisis, can send an <u>inadvertent</u> <u>signal</u> <u>of</u> <u>acquiescence</u> to an adversary, and be misperceived as showing lack of resolve, lack of capability, or a willingness to sacrifice the interests at stake in the crisis in order to avoid an armed clash. The result can be erosion of credibility, undercutting of the nation's bargaining position, and debilitation of efforts to negotiate a solution to the crisis. Even worse, such misperceptions could induce an adversary to preempt in an effort at seeking a military solution with low expectation of concerted resistance. Thus, an inadvertent signal of acquiescence can trigger an inadvertent war of the type begun with a deliberate decision during a crisis.

Efforts to signal resolve or support for an ally, to convey coercive military threats for deterrence or compellence, and to maintain readiness for potential wartime contingencies, can send an <u>inadvertent</u> <u>signal</u> <u>of</u> <u>hostility</u> to an adversary, and be misperceived as showing an intention to seek a military solution to the conflict or to escalate to full-scale war. The result can be an appearance of bad faith which interferes with efforts to negotiate a solution to the crisis, escalation of tensions and hostility in the crisis, and, worst case, a perception by the adversary that war is inevitable leading to a decision by him to preempt rather than suffer the first blow. Thus, an inadvertent signal of hostility can trigger inadvertent war. This can be either the type begun with a preemption decision motivated by the crisis security dilemma--fear of imminent attack--or the type begun with a deliberate decision motivated by the perception that the crisis cannot be satisfactorily resolved short of war.

When a crisis involving the United States arises from a dispute between two other nations, one of whom is an ally or friend of the U.S., the second misperception dilemma comes into play. Studies of naval diplomacy have noted that allies and friends as well as adversaries can misperceive the signals sent by naval forces.[73] Two problems have been

---

[73] Edward N. Luttwak, <u>The Political Uses of Sea Power</u> (Baltimore: Johns Hopkins University Press, 1974), p. 13,

described: first, the danger that signals of support may encourage a friend or ally to be overly aggressive in a conflict, and, second, the danger that reduction in a standing presence, regardless of reason, can be misperceived as signaling reduction in political commitment or even abandonment. These problems comprise the misperception dilemma as it affects signals to allies and friends.

Efforts to signal resolve or support for an ally, to convey coercive military threats for deterrence or compellence, and to maintain readiness for potential wartime contingencies, can send an _inadvertent signal of encouragement_, and be misperceived by the friend or ally as tacit consent for intensification of hostilities and escalation of political demands, or even as overt direct support for initiating fighting which previously had been viewed as infeasible. The result can be an appearance of unlimited commitment which interferes with efforts to negotiate a solution to the crisis, escalation of tensions in the crisis, and outbreak or escalation of fighting in the crisis.

Efforts to signal limited objectives and interest in a negotiated solution, and to limit the level of tension and violence in a crisis, can send an _inadvertent signal of retrenchment_, particularly when the signaling entailed

35; Charles D. Allen, Jr., _The Uses of Navies in Peacetime_ (Washington, DC: American Enterprise Institute, 1980), p. 19.

reduction or withdrawal of a standing presence, and be misperceived by the friend or ally as a signal to restrain his objectives, as wavering commitment or a desire to avoid involvement in the crisis at hand, or even as abandonment. The result can be erosion of credibility with the friend or ally, thus undercutting influence on his behavior, a decision by the friend or ally to seek support from other powers or to build up his military power for autonomous action, or, worst case, a decision by the friend or ally to preempt and seek a _fait_ _accompli_ before his strategic situation worsens further.

Although it would appear logical that an inadvertent signal of hostility to an adversary would tend to be paired with an inadvertent signal of encouragement to an ally or friend, and that an inadvertent signal of acquiescence to an adversary would tend to be paired with an inadvertent signal of retrenchment to an ally or friend, there is no inherent reason for misperceptions to occur in these pairs. Misperceptions result from the individual decisionmaking processes in each nation, responding to stimuli and cognitive factors which can be much different. Thus, while there may be grounds for postulating that certain combinations of perceptions and misperceptions are more likely than others, the occurrence of such combinations should be couched in probabilistic terms rather than described as inherent or inevitable.

In summary, one of the most difficult problems of crisis management is misperception of intentions. The danger of signals sent by military forces being misperceived creates the misperception dilemma. The dilemma in signaling adversaries is between inadvertent signals of acquiescence and inadvertent signals of hostility. The dilemma in signaling allies and friends is between inadvertent signals of encouragement and inadvertent signals of retrenchment. These concepts clarify the problems facing national leaders as they make decisions on employment of military force as a political instrument in crises.

Conclusion

This chapter has reviewed and critiqued the crisis management literature, explaining the major weaknesses in current concepts and presenting new concepts to correct those weaknesses. It began with a review of basic crisis concepts, presenting the perspectives that will be used in the study, followed by a critique of the concept of crisis interaction, particularly the weaknesses in the single interaction sequence model that implicitly underlies existing crisis theories. It then reviewed crisis management concepts, focusing on measures required to maintain control of events in crises, and critiqued the concept of crisis stability, presenting a definition that more accurately reflects the nature of crisis interaction.

Finally, it reviewed the crisis management problem of misperception of intentions and resolve, and presented concepts that more accurately describe the problems decisionmakers face in trying to avoid misperceptions when using military force in crises. The following paragraphs summarize the key new concepts that were presented.

Interaction between the two sides in crises has often in the past been viewed in terms of an implicit single interaction sequence model that does not accurately describe the complexity of crisis interaction. The bureaucratic politics and organizational process models do not entirely correct this weakness and have serious problems of their own. To correct these weaknesses a model containing three interaction sequences--each sequence associated with a specific level in the chain of command--was presented. Crisis interaction is stratified when the military establishment of a country is sufficiently large and complex that national leaders are incapable of exercising direct control over all operational units that could have an impact on a crisis. Paul Bracken's concept of "tightly coupled forces"--modified by the observation that tight coupling is stratified, occurring separately at the strategic and tactical levels without national leaders necessarily being involved--then explains how separate interaction sequences can arise.

One of the most important requirements for the success of crisis management is maintaining close control of

military operations by top-level political authorities. A weakness in the crisis management literature is that it has not adequately addressed the tensions that can arise between crisis management requirements and military considerations.

The crisis security dilemma is that in a crisis, many of the actions a state takes to increase its security and improve its bargaining position decrease the security of the adversary. The most important sources of the dilemma are the military doctrines and the military operations being conducted by the two sides.

The preemption and inadvertent paths to war are particularly relevant to the study of crises. Preemption is motivated by perceptions and fears that the other side is about to strike first. The preemption path to war that results from the crisis security dilemma is the path that will be addressed in this study. Inadvertent war arises from an escalation process in which the two sides employ increasingly threatening military and diplomatic moves in an effort at gaining leverage in crisis bargaining and improving their military positions. Accidents and other inadvertent military actions contribute to the process. The escalation dynamic is driven by rising stakes in the outcome of a conflict, which increase the motivation of national leaders to prevail, and an action-reaction process, in which an escalatory action by one side provokes an escalatory reaction by the other side in recurring cycles.

This escalation dynamic increases tensions and hardens resolve until it results in a deliberate or preemptive decision for war.

One of the most difficult problems of crisis management is misperception of intentions. The danger of signals sent by military forces being misperceived creates the misperception dilemma for national leaders. The dilemma in signaling adversaries is between inadvertent signals of acquiescence and inadvertent signals of hostility. The dilemma in signaling allies and friends is between inadvertent signals of encouragement and inadvertent signals of retrenchment.

The next chapter will build on these concepts to present the theory of stratified interaction and its corollaries of decoupled interactions, the stratified crisis security dilemma, and stratified escalation dynamics.

CHAPTER III

THE THEORY OF STRATIFIED INTERACTION

Studies of crisis management invariably emphasize the importance of top-level political authorities maintaining close control of crisis military operations in order to prevent them from triggering an uncontrollable escalation spiral. Underlying this emphasis on control is concern that interactions between the military forces of the two sides in a crisis could develop their own momentum, decoupled from the political-diplomatic objectives and strategies of national leaders. Although this concern has often been expressed in crisis management studies, the factors that could cause such a decoupling have not been adequately addressed in theories of crisis bargaining and escalation dynamics.

The principle contention of this study is that the single interaction sequense model is inadequate for undertanding the manner in which nations interact in crises, the complexities and difficulties of crisis decisionmaking, and the ways in which crises can get out of control and escalate to war. The theory of stratified interaction developed

117

in this chapter provides a better understanding of these crisis phenomena.

This chapter will begin by describing the stratified interaction model of crisis interactions. With the underlying model in place, the theory of stratified interaction and its first corollary, decoupled interactions, will be defined. The theory of stratified interaction will then be applied to the concept of crisis stability, producing the concepts of the stratified crisis security dilemma and stratified escalation dynamics. Finally, crisis management will be reexamined to show how efforts to prevent stratified interactions from becoming decoupled generate tensions between political objectives and military objectives in a crisis.

The Stratified Interaction Model

The stratified interaction model holds that there are three levels of interaction between the two sides in a crisis: political interaction, strategic interaction, and tactical interaction. These are separate interaction sequences between distinct groups of decisionmakers at each level on both sides in a crisis. In the model, these interactions represent horizontal linkages between decisionmakers at the same level.

Vertical linkages connect decisionmakers at the three levels within each nation. Two types of vertical linkages

connect the three levels: a policy channel and an information channel. There are flows in both directions, downward and upward, in each channel. In the policy channel, there is a flow of orders and policy guidance downward from national leaders to strategic and tactical military commanders, and from strategic level commanders to tactical level commanders. There is also an upward flow of requests for permission to take action, recommended courses of action, reports of intended actions that have not yet been taken, and reports of actions already intitiated that had not been ordered by higher authority. In the information channel, there is a downward flow of intelligence on the adversary, assessments of the adversary's intentions and likely moves, and backgound information on the objectives and strategy being pursued in the crisis. There is also an upward flow of requests for these types of information, intelligence and assessments on the adversary from lower levels, and background information on the situation at the lower levels (such as force readiness data). Vertical interaction between decisionmakers at the three levels in each country takes place through these policy and information channels.

Political interaction is between the top-level political authorities in each nation--the head of government and his immediate advisors, what I have been calling national leaders. In the United States this consists of the

President and those officials present with him during crisis decisionmaking, which normally includes the Secretary of Defense, the Secretary of State, the National Security Advisor, and the Joint Chiefs of Staff (sometimes represented by the Chairman alone). On the Soviet side the political level includes the General Secretary of the Communist Party and certain members of the Politburo, the Defense Council, and the Headquarters (Stavka) of the Supreme High Command (verkhonoye glavnokommandovaniye, VGK, the command element of the Soviet General Staff).[1] Although the term "political" is used to label this top level of interaction, military considerations will, of course, be at least as prominent as political considerations in decision-making. Interactions between the two sides at the political level encompass the full range of diplomatic and military interactions under the cognizance of national leaders.

Strategic interaction is between the strategic nuclear forces and major military commands on each side, thus encompassing conventional as well as nuclear forces. In the United States this includes the Commander in Chief, Strategic Air Command (CINCSAC), the Commander in Chief, Space Command (CINCSPA, which includes the North American

[1] John G. Hines and Phillip A. Petersen, "Changing the Soviet System of Control," International Defense Review 19 (3/1986): 281-289; U.S. Department of Defense, Soviet Military Power, 1986 (Washington, DC: U.S. Government Printing Office, 1986), pp. 12-18.

Air Defense Command, NORAD), the Commander in Chief,

Atlantic Command (CINCLANT, who is also the NATO Supreme

Allied Commander Atlantic, SACLANT), the Commander in Chief,

U.S. European Command (CINCEUR, who is also the NATO Supreme

Allied Commander Europe, SACEUR), the Commander in Chief,

Pacific Command (CINCPAC), the Commander in Chief, Central

Command (CINCCENT), and the Commander in Chief, U.S.

Southern Command (CINCSOUTH, headquartered in Panama). In

the Soviet Union, the strategic interaction level includes

the Strategic Rocket Forces, the National Air Defense Forces

(PVO Strany), and the Commanders in Chief

(glavnokommanduyushchiy) of the Western, Southwestern,

Southern, and Far Eastern Theaters of Strategic Military

Action (teatr voyennykh deystiviy, TVD, often translated as

Theater of Military Operations).[2]

Tactical interaction is between the operational units

(troops, aircraft, and naval vessels) of the two sides.

Tactical interaction occurs primarily at the scene of a

crisis, but can take place anywhere the military forces of

the two sides are operating in close proximity to each

other. Examples, in descending size of the units involved,

would include interaction between the U.S. Sixth Fleet and

the Soviet Mediterranean Eskadra, interaction between a U.S.

---

[2]Ibid. Of the several TVDs identified in Soviet
writings, only the four listed have CINCs appointed to
command them in peacetime.

naval battle group (a carrier or battleship and its escorts) and a Soviet naval task group, and interaction between individual U.S. and Soviet ships or planes. Generally, the small-scale encounters are part of a larger interaction between the military commands on each side responsible fo~ operations in the region encompassing the scene of a crisis. Thus, in the naval realm, tactical interaction will generally be regarded as being between larger units, such as fleets or task forces.

In the stratified interaction model, coupling between the forces of the two sides in a crisis is stratified. Tight coupling at each of the three levels of interaction-- political, strategic, and tactical--can occur because decisionmakers at each level receive direct inputs from warning and intelligence systems. The degreee or "tightness" of coupling at each level can be different, depending on the availability of intelligence and sensors and the strategic and tactical environment (i.e., whether or not tactical-level forces are in close enough proximity for interaction to occur).

The overall U.S. surveillance, intelligence, and early warning system can be viewed as stratified into three levels: national-level assets, strategic warning systems, and tactical sensors. National-level assets include Central Intelligence Agency (CIA) and National Security Agency (NSA) intelligence sources, such as reconnaissance satellites,

electronic and communications intelligence (ELINT and COMINT), human intelligence, and certain reconnaissance missions by military units. Inputs from national-level assets go to decisionmakers at the political level, but much of the intelligence also goes to appropriate lower levels in the military chain of command. The sensitivity and content of the intelligence determine the recipients of it. Distribution of certain intelligence can be restricted to a small group of decisionmakers (and the analysts supporting them), who then make the decision whether or not to promulgate it to lower levels.

Strategic warning systems include the distant early warning (DEW) radar system, early warning satellites, Pave Paws SLBM warning radars, certain ELINT and COMINT systems. Inputs from strategic early warning systems initially go to appropriate decisionmakers at the strategic interaction level. With modern computerized command and control and military data systems, however, certain crucial elements of the information gathered by strategic warning systems can be automatically transmitted to appropriate political level control centers, such as the National Military Command Center and the White House situation room.

Tactical sensors include radar, sonar, visual and photographic reconnaissance, electronic support measures (ESM), and tactical ELINT and COMINT systems. Most inputs from tactical sensors initially go to decisionmakers at the

tactical level, and from there are reported up the chain of command. However, information from certain of the more capable tactical sensor systems is simultaneously reported directly to strategic level commanders, and, with specific prior arrangements, can be transmitted directly to political level control centers.

Two important features these warning and intelligence systems must be noted. First, dual reporting--simultaneous transmission of intelligence to multiple users at various levels in the chain of command--is widely used to expedite the flow of crucial information. Dual reporting generally involves the political and strategic levels, but can also be used with certain tactical sensors. Dual reporting has two effects on the command and control system. On the one hand, it enhances the ability of national leaders to exercise close control of military operations in crises by keeping them better informed of events at the strategic and tactical levels. But, on the other hand, it can increase the autonomy of decisionmakers at the strategic and tactical levels by reducing their dependence on higher authority as a source of warning and intelligence. Which of these two competing tendencies prevails in a particular crisis, or in a particular incident in a crisis, depends upon the specific circumstances in which operational decisions must be made.

The second important feature is that there can be substantial overlap in the coverage of sensors at the three

levels. For example, national-level assets and strategic warning systems can simultaneously detect some military actions. Even tactical sensors can detect some military actions being monitored by strategic warning systems and national-level assets. The effect of overlapping coverage is the same as that of dual reporting: it can enhance top-level control of military operations, or it can increase the autonomy of decisionmakers at the strategic and tactical levels. Overlapping coverage thus can either intensify or inhibit stratification of crisis interaction.

Figure 1 provides a diagram of the stratified interaction model. The circles represent decisionmakers on the two sides: P designates political level decisionmakers, S designates strategic level decisionmakers, and T designates tactical level decisionmakers, with the subscripts designating the two sides. The diamonds represent interactions between the two sides: I_p is political interaction, I_s is strategic interaction, and I_t is tactical interaction. Horizontal arrows from circles to diamonds represent actions the two sides take toward each other, horizontal arrows from diamonds to circles represent detection of the other side's actions.[3] Vertical arrows represent flows of information (upward and downward), orders

---

[3]This is a simplification of the interaction loop used by Paul Bracken, The Command and Control of Nuclear Forces (New Haven, CT: Yale University Press, 1983), p. 62.

(downward), and requests for permission to take specific
actions or recommendations that specific actions be taken
(upward). The P_1-T_1 arrow and the P_2-T_2 arrow represent
efforts by political level decisionmakers to exercise
direct, positive control of operational forces in a crisis.

Figure 1. The Stratified Interaction Model
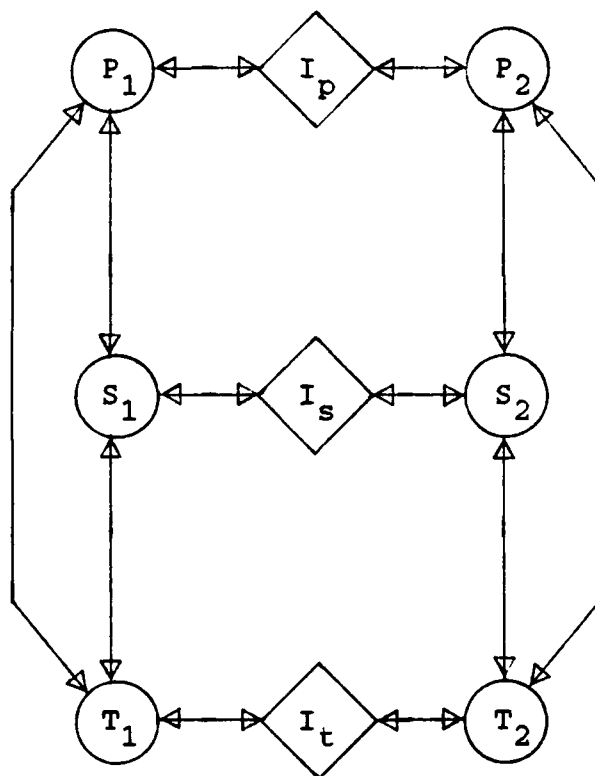
In summary, the stratified interaction model states
that there are three levels of interaction between the two
sides in a crisis: political interaction, strategic
interaction, and tactical interaction. Political
interaction is between the top-level political authorities
in each nation. Strategic interaction is between the

strategic nuclear forces and major military commands on each side, encompassing conventional as well as nuclear forces. Tactical interaction is between those operational military units (troops, aircraft, and naval vessels) of the two sides that are in direct contact. Separate interaction sequences at these three levels is possible because military commanders are delegated significant deicionmaking authority and receive direct inputs from warning and intelligence systems on the adversary's military activies . In addition to these three horizontal interaction sequences between the two sides, there is also vertical interaction between decisionmakers at the three levels in each country. These vertical interactions take place through the policy and information channels that link the three levels.

The Theory of Stratified Interaction

The theory of stratified interaction can now be stated: Given conditions of delegated command, tight coupling, and acute crisis, interactions between the two sides will have a tendancy to become stratified into separate political, strategic and tactical interactions. The definitions of the political, strategic, and tactical levels of interaction are as given above for the stratified interaction model.

As stated in the definition of the theory, three conditions contribute to stratified interaction. First, the

military establishments of the two sides are sufficiently
large and complex that top-level political authorities
cannot exercise constant, direct, positive control over the
actions of all operational units, and must therefore rely to
a large degree on delegated command. Second, the military
forces of the two sides are tightly coupled through warning
and intelligence systems that are vertically integrated with
major military commands and operational forces. This
condition is driven by perceptions that striking first will
accrue significant strategic or tactical advantages, thus
requiring warning of attack to ensure the survival of
operational forces and the ability to effectively execute
wartime contingency plans. Third, stratified interaction
occurs in an acute international crisis, when military
forces are being used as political instrument for crisis
bargaining. This results in actions being taken with
military forces that deliberately or inadvertently convey a
military threat to the other side.

Strategic and tactical level interactions also occur
under normal peacetime conditions and in lesser crises that
do not pose a danger of war. In fact, under normal
peacetime conditions, when national leaders are paying very
little attenion to routine military operations, there could
be numerous interaction sequences taking place between
forces in direct contact with the other side's forces. But
such peacetime interactions are normally not of great

interest because they do not gather momentum or seriously affect the overall state of relations between the two sides. This is because the perception of an acute danger of war is not present and the interactions do not occur in the context of deliberate efforts to convey military threats to the other side.

An example of this is peacetime incidents at sea between U.S. and Soviet naval forces. The U.S. and Soviet fleets are almost always in close proximity somewhere in the world and their interactions occasionally produce incidents, such as shouldering (forcing a ship clear of a formation), threatening actions with weapons, and even collisions. But such incidents have never produced more than diplomatic protests, even at the height of the cold war before the 1972 U.S.-Soviet Incidents at Sea Agreement was signed. Another example is the dozens of American military aircraft fired on or shot down by the Soviet Union, China, and the Warsaw Pact countries during the 1950s and 1960s.[4] None of these Cold War incidents resulted in tactical level interactions between the forces of the two sides that gained their own local momentum, even though U.S. leaders responded to a few of the incidents with military shows of force.

The existence of stratified interaction in a crisis is not in itself important. The interaction sequences at the

---

[4]See Chapter V for a detailed discussion of these incidents.

strategic and tactical levels can be expected to parallel the interaction sequence at the political level so long as national leaders are able to control the overall magnitude and momentum of military operations--even if they cannot control every operational decision in the interaction sequence. The three interaction sequences are described as being parallel when the intensity of the hostilities and magnitude of threat (or reassurance) being signaled by strategic and tactical level interactions are roughly what national leaders desire to implement their political-diplomatic strategy for managing the crisis.

What is of analytical interest is the decoupling of interactions at the three levels, which could cause national leaders to lose control of events in a crisis and touch off an escalatory spiral. Decoupled interactions are defined to be an interaction sequence at the strategic or tactical level in which the intensity of hostilities, level of violence, and magnitude of threat being conveyed to the other side are not under the control of national leaders. This can occur when there is some sort of interruption or severe degradation of the vertical policy and information channels between decisionmakers at the three levels.

When interactions are decoupled, the three interaction sequences are no longer parallel. The intensity of hostilities at the strategic or tactical levels no longer supports the political-diplomatic strategy being prusued by

national leaders in the crisis. In principle, decoupling
can lead to the intensity of hostilities at the strategic
and tactical levels being either greater or lesser than that
desired by national leaders. Although the escalation
dynamics theory predicts that the tendency would normally be
toward escalation of hostilities, this is a question for
empirical research.

The first corollary to the theory of stratified
interaction is that decoupling of interactions will occur to
the extent that operational decisions on the employment of
military forces made at the strategic and tactical levels
differ from the operational decisions political level
decisionmakers would have made to coordinate those military
actions with their political-diplomatic strategy for
resolving the crisis. Conversely, decoupling of stratified
interaction is averted to the extent that political level
decisionmakers exercise constant, direct, positive control
over operational military forces, or ensure that the
guidance contained in mechanisms of delegated command
produce operational decisions at the strategic and tactical
levels that support their political-diplomatic strategy for
resolving the crisis.

This is not to imply that national leaders always act
wisely while exercising direct control over their military
forces, or that tactical-level military commanders have a
propensity to disrupt crisis management efforts when not

under direct top-level control. Inept or indiscriminate employment of military force by national leaders can defeat crisis management efforts as easily as inappropriate operational decisions by on-scene commanders. Additionally, as will be seen in the case studies, on-scene commanders are quite able to act with prudence and caution when not under direct control by national leaders. This is discussed in detail in Chapter IX.

Military commanders are never without operational guidance of some sort. When direct control is interrupted, for whatever reason, they will base operational decisions on the last direct guidance received until the tactical circumstances change sufficiently to make that guidance inapplicable--which can happen very quickly. At that point they revert to the operational guidance contained in the mechanisms of delegated command. This is an entirely rational system of command, reflecting the reality that decisionmaking at the tactical level does not cease simply because national leaders are unable to make the decisions. Once operational forces make this de facto shift to delegated command, even though nominally still under direct command, decoupling of interactions can occur.

There are seven potential causes of decoupling: communications and information flow problems, impairment of political level decisionmaking, a fast-paced tactical environment, ambiguous or ambivalent orders, tactically

inappropriate orders, inappropriate guidance in mechanisms of indirect control, and deliberate unauthorized actions by military commanders. As these potential causes suggest, decoupling can occur even when national leaders are attempting to exercise constant, direct, positive control over operational forces.

Communications problems can sever the links from national leaders to operational forces, leaving those forces at least temporarily under delegated command. The problems can take many forms, including outright loss of radio contact, garbled messages, delays in message delivery due to system overload, misrouting of messages, and deliberate interference by the adversary. Although the U.S. military communications system has been vastly improved over the last four decades--without which the President could not even attempt to exercise close control of military operations--it is still not infallible.

A wide range of information problems can contribute to decoupling. Information flows can be interrupted by communications problems, excessive secrecy and compartmentation, or even a simple failure to realize that a particular report warrants the immediate attention of decisionmakers. Exclusive information is a resource that confers influence on policy decisions, which can lead to hoarding or hiding of crucial facts. Too much information can also cause problems, particularly when large ammounts of inaccurate and

irrelevant information must be sifted out to reveal what is accurate and relevant. Accurate assessment and effective use of information can be degraded by the cognitive limits on analysis and decisionmaking.[5] Information problems such as these can prevent national leaders from exercising effective direct control over crisis military operations by leaving them unaware of the need to make certain operational decisions, denying them the capability to make those decisions, or convincing them that they should delegate the decisions to lower levels.

Impairment of top-level decisionmaking under the stress of a crisis, or preoccupation with a particular aspect of a crisis, can result in real-time guidance not being provided to operational forces even when communications channels are intact. Selective and sequential attention to problems is a well-recognized cognitive limit on decisionmaking. When decisionmakers become overloaded with information and urgent problems, "load shedding," to use Coral Bell's apt expression, can occur, producing

---

[5]See Roberta Wholstetter, <u>Pearl Harbor: Warning and Decision</u> (Stanford, CA: Stanford University Press, 1962), p. 394; Harold L. Wilensky, <u>Organizational Intelligence: Knowledge and Policy in Government and Industry</u> (New York: Basic Books, 1967), Chapter 3; Anthony Downs, <u>Inside Bureaucracy</u> (Boston: Little, Brown and Co., 1967), p. 75; Ole R. Holsti, <u>Crisis Escalation War</u> (Montreal: McGill-Queen's University Press, 1972), pp. 104-118; Alexander L. George, <u>Presidential Decisionmaking in Foreign Policy</u> (Boulder CO: Westview Press, 1980), pp. 121-36; Ronald H. Hinkly, "National Security in the Information Age," <u>The Washington Quarterly</u> 9 (Spring 1986): 125-40.

inaction or much delayed reactions when a new problem arises.[6] Such impairment of decisionmaking by national leaders can rob them of effective direct control of military operations, resulting in tactical or strategic level interactions being decoupled from political level guidance.

In a fast-paced tactical environment, tactical decisionmakers may not have time to describe their circumstances to national leaders and await a decision before having to take action. To use an exaggerated example, when missiles are inbound the captain of a ship cannot wait for the National Security Council to convene and haggle over his fate. For this reason, operational commanders always have a certain amount of decisionmaking authority delegated to them, regulated by the rules of engagement. When urgent operational decisions must be made on the basis of delegated command rather than on consultation with national leaders, what might be called momentary decoupling occurs. If direct command is immediately reimposed, decoupling ceases. But if an action-reaction sequence starts at the tactical level,

---

[6]Coral Bell The Conventions of Crisis (London: Oxford University Press, 1971), pp. 88-89. Also see the discussions of sequential attention to goals and problemistic search in James G. March and Herbert A. Simon, Organizations (New York: John Wiley and Sons, 1958), Chapter 6; Richard M. Cyert and James G. March, A Behavioral Theory of the Firm (Englewood Cliffs, NJ: Prentice-Hall, 1963), pp. 118-122; John D. Steinbruner, The Cybernetic Theory of Decision: New Dimensions of Political Analysis (Princeton, NJ: Princeton University Press, 1974), pp. 72-74; George, Presidential Decisionmaking, pp. 25-53.

with operational commanders on both sides making decisions on the basis of delegated command, momentary decoupling could lead to an interaction sequence that is decoupled from the political-diplomatic strategy of national leaders.

National leaders, uncertain as to the implications of their political-diplomatic strategy for tactical military operations or even uncertain as to the strategy itself, may issue orders to military forces that are ambiguous or ambivalent. To some degreee this problem is inherent in the nature of crisis management, which consists of the dual goals of protecting vital national interests while avoiding unwanted escalation of the confrontation. Ambiguous or ambivalent orders are particularly likely when they must be formulated under the stress and time pressures of a crisis, Thorough evaluation of alternative tactical options may not be possible before an order must be given. A military commander faced with ambiguous or ambivalent orders may not have time to seek guidance on how to interpret them in a specific situation, forcing him to rely on his own best judgement. If his decisions, no matter how carefully reasoned or tactically appropriate they may be, do not support the political-diplomatic initiatives being pursued to manage the crisis, then decoupling has occurred. If top-level control can be immediately re-established, the decoupling will only be momentary. But if it cannot, the momentary decoupling could lead to an interaction sequence

that is decoupled from the political-diplomatic strategy of national leaders.

National leaders could well decide to exercise only a small degree of direct control over certain military operations during a crisis, relying instead on military commanders to carry out their wishes. When this occurs, the guidance contained in mechanisms of indirect control--the alert system, standing orders, mission orders, contingency plans, and rules of engagement--becomes crucial to effective crisis management. If national leaders do not pay sufficient attention to that guidance, military actions could occur that they had not anticipated and which exceed the scope of operations they had desired.[7] This could cause momentary decoupling and lead to an interaction sequence that is decoupled from the political-diplomatic strategy being pursued by national leaders.

The final possible cause of decoupling is a deliberate unauthorized action by a military commander. In this case the commander has specific orders for the mission he is to carry out, but knowingly decides to disobey those orders and carry out an action contrary to the letter and intent of his orders. A military commander might do this because he

---

[7]Scott D. Sagan has suggested that this was the case with the scope of U.S. Navy anti-submarine warfare operations during the Cuban Missile Crisis. "Nuclear Alerts and Crisis Management," International Security 9 (Spring 1985): 117-118.

disagrees with the political-diplomatic strategy being
pursued by national leaders and seeks to achieve what he
considers to be a superior resolution of the crisis. More
likely, however, would be the case in which a military
commander deliberately takes an unauthorized action because
he perceives his local tactical situation as being much more
threatening than do national leaders, or believes that his
orders are infeasible under the conditions he faces. The
military commander's assessment of the situation could be
entirely correct, but the action is still unauthorized.

In summary, the theory of stratified interaction
states that, given conditions of delegated command, tight
coupling, and acute crisis, interactions between the two
sides will become stratified into separate political,
strategic and tactical interactions. A corollary to the
theory is that decoupling of stratified interactions will
occur to the extent that operational decisions on the
employment of military forces made at the strategic and
tactical levels differ from the operational decisions
political level decisionmakers would have made to coordinate
those military actions with their political-diplomatic
strategy for resolving the crisis.

Stratified Interaction and Crisis Stability

The concept of the crisis security dilemma, as defined
in the previous chapter, is that in a crisis, many of the

actions a state takes to increase its security and improve
its bargaining position decrease the security of its
adversary. Applying the theory of stratified interaction to
the concept of crisis stability produces the second
corollary to the theory, the stratified crisis security
dilemma: In an acute crisis, the security dilemma is
stratified, arising from the interaction processes occurring
separately at each of the three levels, and affecting the
likelihood of war separately at each level.

This corollary contends that the adversary's military
intentions--whether they are essentially offensive or
defensive--can be perceived differently by decisionmakers at
the political, strategic, and tactical levels of crisis
interaction. Many military moves are ambiguous as to their
offensive or defensive intent, and can increase capabilities
in both areas. Interactions at the strategic and tactical
levels can generate circumstances in which actions taken by
one side to increase the security of their forces or improve
their tactical position can decrease, or appear to decrease,
the security of the other side's forces. Actions by one
side prompt countermeasures by the other side that in turn
contribute to an escalatory action-reaction spiral as
military commanders on both sides seek to maintain or
increase their strategic or tactical advantages.

The danger at the strategic and tactical levels is
that at some point in the interaction one side will take an

action that increases the perceived threat of attack against the other side to an intolerable level, prompting the adversary to preempt. Military commanders could have authority to preempt for self-defense under conditions specified in their standing orders or rules of engagement, or could be required to seek authority to preempt from national leaders. The key point is that the interaction process that created the circumstances in which preemption was perceived to be necessary was not under control of national leaders.

The mutually reinforcing alert phenomenon described by Bracken is an example of the stratified security dilemma at the strategic level of interaction.[8] An example at the tactical level of interaction would evolve like this: movements of naval forces intended to signal resolve by placing them within striking range of the adversary's naval forces increase the vulnerability of the adversary's forces, prompting them to take measures--authorized in their standing orders--to increase their ability to defend themselves. Those defensive measures, in turn, increase the

---

[8]Bracken, pp. 64-65. In this example, actions taken by the commanders of one side's strategic forces (in accordance with their standing orders) intended to reduce the vulnerability of those forces are quickly detected by the commanders of the other side's strategic forces, who cannot distinguish those actions from preparations for offensive action, and must therefore take actions to decrease the vulnerability and increase the readiness of their own forces.

vulnerability of the first side's naval forces, prompting them to take measures--once again, authorized in standing orders--to increase their ability to defend themselves.

This type of interaction occurred between U.S. carrier battle groups and Soviet anti-carrier forces in the Mediterranean during the 1973 Middle East War. Soviet ships and submarines armed with anti-ship cruise missiles moved into positions where they could launch preemptive strikes against the U.S. Sixth Fleet on short notice. This in turn prompted the Sixth Fleet to maneuver to evade being targeted, and to deploy ships, submarines, and armed aircraft into positions where they could strike Soviet cruise missile platforms upon indication of an attack.[9] This maneuvering for tactical advantage continued throughout the crisis until U.S.-Soviet tensions subsided and national leaders on both sides ordered their naval forces in the Mediterranean to standdown and resume peacetime operations.

---

[9]Lieutenant F.C. Miller, "Those Storm-beaten Ships, Upon Which the Arab Armies Never Looked," U.S. Naval Institute Proceedings 101 (March 1975): 23-24; Admiral Elmo R. Zumwalt, Jr., On Watch: A Memoir (New York: Quadrangle, 1976), pp. 436-47; Robert G. Weinland, "Superpower Naval Diplomacy in the October 1973 Arab-Israeli War: A Case Study," in The Washington Papers, vol. 6, no. 61 (Beverly Hills: Sage, 1979), pp. 68-88; Charles D. Allen, Jr., The Uses of Navies in Peacetime (Washington, DC: American Enterprise Institute for Public Policy Research, 1980), pp. 30-33. Also see Admiral Isaac C. Kidd, Jr., "View From the Bridge of the Sixth Fleet Flagship," U.S. Naval Institute Proceedings 98 (February 1972): 18-29; Admiral Stansfield Turner and Commander George Thibault, "Countering the Soviet Threat in the Mediterranean," U.S. Naval Institute Proceedings 103 (July 1977): 25-32.

The stratified crisis security dilemma provides the basis for the third corollary to the theory of stratified interaction, <u>stratified escalation dynamics</u>: In an acute crisis, in which strategic or tactical interactions between the two sides have become decoupled from political level interactions (meaning that the strategic or tactical interactions are no longer under the direct or indirect control of national leaders), the security dilemma, operating separately at each level, can trigger an escalatory spiral at the strategic or tactical levels of interaction, which under certain circumstances can cause the crisis to escalate uncontrollably to war.[10] An escalation spiral can be touched off at any of the three levels. If it starts at the political level, with national leaders making the escalatory decisions, it immediately encompass all three levels and thus is not stratif: J.

If an escalation spiral starts at the tactical or strategic level, it will not necessarily be transmitted upward to higher levels of interaction. National leaders could, for example, decide to let an uncontrollable escalation spiral between their forces in direct contact in the field or at sea (the tactical level) play itself out, and decide not to escalate the war at the strategic level.

---

[10]On how such escalation processes work, see Phil Williams, <u>Crisis Management</u> (New York: John Wiley, 1976), pp. 97, 101.

constant, direct, positive control of operational forces at the strategic and tactical level, or national leaders ensuring that the operational guidance in the mechanisms of indirect control supports their political-diplomatic strategy for resolving the crisis. Both of these means of controlling escalation can be difficult to implement in practice. As was dicussed earlier, direct control can be interrupted or degraded by a wide range of technical, operational and even psychological impediments. Ensuring that the mechanisms of indirect control support the political-diplomatic strategy in a crisis is made exceedingly difficult by the near impossibility of anticipating every possible tactical situation that a military commander might face. Thus, there are ample grounds for postulating that an escalation spiral beyond the control of national leaders could arise at the tactical or strategic levels.

A key point of the stratified escalation dynamics concept is that the phenomenon can exist under an assumption of rational decisionmaking. Fully rational strategic and tactical level decisionmakers, acting strictly in accordance with approved operational guidance, could well initiate an action-reaction sequence that becomes an uncontrollable escalation spiral. There is no need to assume military evasion of civilian control, as is done in one definition of inadvertent escalation, or crazed officers disobeying

orders, which is one of the accidental war scenarios.[11]
Both are highly unlikely and the record of U.S. and Soviet
behavior in superpower crises shows a great deal of caution
on the part of operational commanders, rather than excessive
aggressiveness.

To contend that stratified escalation dynamics can
occur under an assumption of rational decisionmaking is not
to deny that the stress and confusion of a crisis can
exacerbate the cognitive constraints on decisionmaking.
Misperception and miscalculation are highly likely during a
crisis, and increase the possibility that crisis military
operations could touch off an escalation spiral. The crisis
security dilemma creates a decisionmaking environment in
which misperception and miscalculation are not only more
likely to occur, but are also more likely to touch off an
escalation spiral.

Stratified interactions can cause misperceptions of
the intentions of adversaries. Strategic and tactical level
interactions that result from military commanders acting on
their own initiative, responding within the bounds of their

---

[11]Military evasion of civilian control is a central
element of Barry R. Posen's definition of inadvertent
escalation. See his "Inadvertent Nuclear War? Escalation
and NATO's Northern Flank," International Security 7 (Fall
1982): 31-32. For a description of the accidental war
thesis see Paul Bracken, "Accidental War," in Graham T.
Allison, Albert Carnesale, and Joseph S. Nye, Jr., eds.,
Hawk Doves & Owls: An Agenda for Avoiding Nuclear War (New
York: W.W. Norton, 1985), pp. 25-53.

authority to actions by the other side, can be misperceived as having been deliberately instigated. If viewed as having been ordered by national leaders, strategic and tactical level interactions are interpreted in the overall political and military context of the crisis, rather than the more limited immediate circumstances that surrounded the interaction. This can result in strategic and tactical level interactions being viewed as signals of the intentions of the adversary's leaders, which may not have been the case at all. Misperception of intentions can thus arise from interactions among forces at the strategic and tactical levels as well as from the signals being exchanged at the political level.

Inadvertent military incidents are most likely to occur under conditions of stratified interaction, when national leaders are relying on mechanisms of delegated command for indirect control of military forces. An inadvertent military incident could spark an action-reaction escalation sequence at the tactical or strategic level decoupled from interaction at the political level as military commanders acted in accordance with mechanism of delegated command. If national leaders then misperceive the escalating military engagement as a deliberate signal of hostile intent or as a direct military threat to their security, a situation likely under the conditions of the crisis security dilemma, stratified escalation dynamics

could spread upward, affecting all three levels of interaction and leading to war.[12]

Serious fighting could also erupt without escalation dynamics spreading to the political level, that is, while national leaders on the two sides were still trying to resolve the crisis without war. An intense engagement between the forces of the two sides in direct contact in the field or at sea (the tactical level) could spread upward to the major theater commands in charge of those forces (the strategic level) through the operation of delegated decision-making authority at the strategic level. Actions taken by strategic-level military commanders on both sides for essentially defensive purposes, such as increasing the readiness of conventional and nuclear forces and initiating measures to support the forces already engaged at the tactical level, could set in motion an escalating action-reaction cycle at the strategic level decoupled from the objectives of national leaders.[13] The most likely scenario

---

[12]The crisis management technique of deliberately slowing the tempo of military operations and creating pauses for exchange of diplomatic communications is intended to halt such an escalation process. See Alexander L. George, "The Development of Doctrine and Strategy," in Alexander L. George, David K. Hall and William E. Simons, The Limits of Coercive Diplomacy (Boston: Little, Brown and Co., 1971), p. 9; and "Crisis Management: The Interaction of Political and Military Considerations," Survival 26 (September/October 1984): 226.

[13]The mutually reinforcing alert described by Paul Bracken is an extreme example of this. See Bracken, Command and Control of Nuclear Forces, pp. 64-65.

for war under these conditions would be for conventional forces in the field, brought to a wartime readiness posture by escalation dynamics at the strategic level, to react to deliberate or inadvertent actions taken by the adversary's forces--actions allowing a defensive response under their rules of engagement--by initiating combat actions in accordance with wartime contingency plans.

In summary, the stratified crisis security dilemma is that, in an acute crisis, the security dilemma is strati-fied, arising from the interaction processes occurring separately at each of the three levels, and affecting the likelihood of war separately at each level. This leads to the the stratified escalation dynamics corollary: in an acute crisis, in which interaction between the two sides has become stratified and decoupled, the security dilemma, operating separately at each level, can trigger an escalatory spiral at the strategic or tactical levels of interaction, which under certain circumstances can cause the crisis to escalate uncontrollably to war.

Political-Military Tensions

Awareness of the problems inherent in crisis management and the escalatory impact they could have on a crisis underlies the emphasis national leaders place on maintaining close control of military forces in crises. An irony of crisis management is that efforts to prevent

interactions at the strategic and tactical levels from
becoming decoupled from the political-diplomatic strategy
being pursued to manage a crisis can generate tensions
between political and military considerations that create
further difficulties for managing the crisis. These
tensions and the manner in which they are resolved directly
affect, and are directly affected by, stratified
interaction. Tensions between political and military
considerations are inherent in the use of military force as
a political instrument. They would arise in a crisis even
without stratified interaction, but their impact on the
ability of national leaders to manage a crisis are
exacerbated under conditions of stratified interaction.

These tensions are actual or potential conflicts
between political and military considerations which force
decisionmakers, knowingly or tacitly, to make trade-offs
among individually important but mutually incompatible
considerations.[14] Three such tensions arise in crises:
crisis objectives tensions, operational control tensions,
and wartime readiness tensions.

The first source of tension is conflicts between
political and military considerations: tensions between
political considerations and the needs of diplomatic

---

[14]The concept of tensions between political and
military considerations is derived from Alexander L.
George's concept of "interaction of political and military
considerations." George, "Crisis Management," pp. 223-234.

bargaining, on the one hand, and military considerations and the needs of military operations, on the other. Both sets of considerations are those being pursued to influence the outcome of the crisis. Tensions between political and military considerations arise because military forces must always be prepared for the possibility of combat even when being used for political signaling. Thus, such tensions arise well before force is actually used in a crisis.

There are two approaches to using military forces deployed to the scene of a crisis for political signaling. In the first approach—the symbolic or indirect threat— token forces are sent to signal resolve to protect vital national interests and as a symbol of the overall military power of the nation. Token deployments are often used in conjunction with increases in the readiness of other forces in the theater or forces held in strategic reserve for rapid deployment to the scene. The token forces at the scene convey a threat that other, more powerful forces will be used if the target nation does not respond in a satisfactory manner to the threat being conveyed. In some circumstances token forces serve as a tripwire, enhancing, through the likelihood of their being involved in any fighting that erupts, the credibility of a deterrent threat.[15]

---

[15]Thomas C. Schelling, <u>Arms and Influence</u> (New Haven, Conn.: Yale University Press, 1966), p. 47; Henry A. Kissinger, <u>Nuclear Weapons and Foreign Policy</u> (New York: Harper and Brothers, 1957), p. 242; Glenn H. Snyder,

Tension between political and military considerations can arise even when military forces are used in this limited manner. The token forces typically do not have a mission other than to be present at the scene of the crisis, leaving them without a clear military objective other than survival, and even that may be compromised when the forces are placed in an exposed position as a tripwire. Military moves by the adversary that are hardly worth the attention of national leaders can present an imminent threat to the token force. It is thus to be expected that on-scene military commanders will have perceptions and priorities much different from those of national leaders in this situation.

In the second approach to using military forces for political signaling, strong forces (well beyond what would be needed for a token force) are used to convey a direct threat with their intrinsic warfighting capabilities. Alerting strategic nuclear forces during a crisis is an example of a direct threat, intended to achieve deterrence by threat of punishment. Deployment of substantial ground, air, or naval forces capable of engaging the adversary's forces at the scene of a crisis is another example of a direct threat, intended to achieve deterrence by threat of denial.

Deterrence and Defense (Princeton, NJ: Princeton University Press, 1961), pp. 130-131; Lawrence Freedman, The Evolution of Nuclear Strategy (New York: St. Martins's, 1983), pp. 90, 290-291.

When military forces are used in a crisis to convey a
direct threat, the tension between political and military
considerations can become acute. This is because the
credibility of the threat being conveyed by the forces is a
function of the adversary's perception of their capability
to carry out the military actions being threatened, as well
as the credibility of the threat to use them if necessary.
The adversary's perception of the likelihood of the forces
actually being used is certainly the more important
consideration, but his perception of their capabilities can
influence that assessment. For example, a threat to
intervene with a battalion of troops to halt an invasion by
a division-sized force would probably not be credible.
Additionally, three separate groups of adversary decision-
makers--at the political, strategic, and tactical levels--
are assessing the credibility of the force. Different
decisionmakers could well focus on different indicators of
intent. Military commanders, particularly those commanding
forces in contact with the adversary force, are likely to be
more attentive to capabilities as an indicator of intentions
than are national leaders weighing a broader range of
considerations.

The capability of forces at the scene of a crisis to
carry out the military actions being threatened is a
function of three factors: strength, readiness, and tactical
situation. Strength, the material dimension of credibility,

is the warfighting capability of the forces employed
relative to the adversary forces they would have to fight.
It is a product of quantity of forces and specific types of
forces and weapons weapons employed. Readiness and tactical
situation are the operational dimension of credibility.
Readiness refers to the readiness posture of the forces
employed: the degree to which they are prepared to conduct
combat operations. It is a product of manpower, logistics
(fuel and ammunition for combat), and the operational
procedures in effect (arming of aircraft on patrol, manning
of weapons on ships, or deploying troops in combat units).
Tactical situation refers to the impact of geographic
position, relative to adversary forces, on the the ability
of the forces employed in a crisis to effectively carry out
combat operations. The degree to which a particular
tactical situation is advantageous (and therefore credible)
or disadvantageous (and therefore not credible) is a complex
calculation involving geography, the capabilities of each
side's weapons, surveillance and warning capabilities, and
the speed required by each side to react to threats and
launch attacks.

The nature of the tensions between political and
military considerations will be illustrated by contrasting
military considerations with three of the requirements for
crisis management identified by Alexander George: (a)
deliberately slowing the tempo of military (erations in

order to create pauses for the exchange of diplomatic signals, assessment, and decisionmaking, (b) coordinating military actions with diplomatic actions in an integrated strategy for resolving the crisis acceptably without war, and (c) avoiding military options that give the adversary the impression of an impending resort to large-scale warfare.[16] These crisis management requirements can have two effects on the military forces deployed to the scene of a crisis: they can adversely affect the three elements of warfighting capability--strength, readiness, and tactical situation--and they can severely complicate the tactical planning of military commanders.

The elements of warfighting capability (strength, readiness, and tactical situation) were described above. The second military consideration affected by crisis management, tactical planning, requires elaboration. In formulating tactical plans for the conduct of combat operations, military commanders seek to adhere to or exploit certain operational considerations believed to provide tactical advantages in combat, commonly referred to as the principles of war.[17] Four of these principles are

---

[16]George, "Crisis Management," p. 226.

[17]Arguments over the value and proper formulation of the principles are irrelevent to the point being made. Regardless of their specific wording or ranking, the principles provide a convenient describe of what military commanders try to do in battle. On the origins of the principles, see John I. Alger, The Quest for Victory

especially affected by crisis management requirements:
objective, initiative, concentration, and surprise.

The principle of the objective is that military
operations must be directed toward a clearly defined,
decisive, and attainable military objective. The principle
of initiative, often called the principle of the offensive,
is that seizing the initiative with offensive action is
almost always necessary to achieve decisive results, to
maintain freedom of action (choosing when and where to
engage the enemy, so as to exploit his weaknesses), and to
control the pace and course of battle. Exploitation (or
pursuit)--rapidly following up initial success with further
offensive action--is an element of initiative, intended to
keep the enemy off balance and on the defensive. The
principle of concentration, often called the principle of
mass, is that superior combat power, a function of quality
as well as quantity of arms, must be concentrated at the
critical time and place for a decisive purpose. The
principle of surprise is that striking the enemy at an
unexpected time and place, and in a manner for which he is
not prepared, can decisively shift the balance of combat

(Westport, CT: Greenwood Press, 1982). For a critique of
the principles see Bernard Brodie, "Strategy as a Science,"
World Politics 1 (July 1949): 467-488. For contemporary
applications of the principles see Colonel Harry G. Summers,
On Strategy: The Vietnam War in Context (Carlisle Barracks,
PA: U.S. Army War College, April 1981), pp. 53-100; Major
Robert L. Earl, "A Matter of Principle," U.S. Naval
Institute Proceedings 109 (February 1983): 29-36.

power, achieving much greater success for the effort
expended.[18]

Having defined the elements of warfighting capability
(strength, readiness, and tactical situation) and selected
principles of tactical planning (objective, initiative,
concentration, and surprise), we can now examine why
tensions arise between crisis management requirements and
military considerations. For brevity, this discussion will
present only one aspect of the tensions: the impact of
crisis management requirements on military considerations.
The reverse aspect--the impact of military considerations on
crisis management requirements--should be obvious in each
case. Essentially, unrestrained pursuit of the military
principles will usually preclude meeting the crisis
management requirements.

Deliberately slowing the tempo of military operations
violates the principle of initiative, particularly its
component principle of rapidly exploiting initial success
with further offensive action, and makes it difficult to use
surprise, which depends in part on speed of execution. By
not allowing one's forces to seize the initiative, thus
granting the adversary's forces at least partial control

---

[18]Definitions are taken from U.S. Department of the
Army, Field Manual 100-5, Field Service Regulations:
Operations (Washington, DC: U.S. Government Printing Office,
19 February 1962), pp. 46-48, but are simplified and reflect
some terms used by Summers and Alger.

over the tactical situation, deliberately slowing the tempo
of military operations can place military forces in an
unfavorable tactical situation.

Coordinating military actions with diplomatic actions
can have a wide range of negative impacts. Coordination is
sought to support an integrated strategy for resolving the
crisis acceptably without war, often with the result that
the military objectives of crisis operations are not well
defined--violating the principle of the objective. In
military planning, objectives are specific and concrete:
such as defending or seizing a specific location, or
defeating a specific force. A principle objective of
crisis management is to avoid war or unwanted escalation of
a limited use of force, which is difficult for military
planners to use as the basis for planning tactical
operations. Resolving the crisis acceptably is achieved
through political-diplomatic bargaining, in which military
forces are used primarily as a means of signaling intentions
and conveying coercive threats, a political mission
difficult for military planners to address as an objective
in the military sense.

Coordinating military actions with diplomatic actions
can also require limiting the size and composition of the
forces employed, violating the principle of concentration
and constraining the strength of the force, an element of
its warfighting capability. Political signaling can require

deploying military forces close to or in the midst of
fighting, or close to the adversary's forces, as a visible
signal of commitment, thereby placing one's forces in an
unfavorable tactical situation. Limiting the actions that
military forces can take in self-defense or to improve their
tactical situation, normally done for escalation control
purposes, can violate the principle of initiative and
constrain their readiness for combat. Informing the
adversary of one's military operations, an action normally
taken to reinforce the signal being conveyed by military
forces but which can also serve escalation control purposes,
deliberately violates the principle of surprise and
secondarily violates the principle of initiative (It is hard
gain control of the tactical situation when the adversary
knows what you are doing and why you are doing it). Using
force in gradually increasing increments, a common tactic of
coercive bargaining, is the military planner's second worst
nightmare (second only to being the victim of strategic
surprise), violating the principles of initiative
(particularly exploitation), concentration, and surprise.

Avoiding military options that give the adversary the
impression of an impending resort to large-scale warfare can
also have a broad range of impacts on military considera-
tions. This crisis management requirement, which is
essentially application of the escalation avoidance strategy
under crisis conditions, can require limiting the size and

composition of the forces employed, limiting the tactical
actions they are permitted to take, and using force in
gradually increasing increments. These approaches to the
use of military force violate all four of the principles of
tactical planning (objective, initiative, concentration, and
surprise), and constrain the elements of warfighting
capability (strength, readiness, and tactical situation).
The principles of tactical planning essentially dictate that
the military options favored by military planners will be
precisely those which give the adversary the impression of
an impending resort to large-scale warfare, or the appear-
ance of an actual resort to large-scale warfare in the case
of surprise attack. An optimum battle plan calls for
seizing the initiative by concentration of superior force
and launching a surprise attack against a strategic objec-
tive, and then rapidly exploiting that attack with further
offensive action. Such a battle plan is precluded by the
requirements of crisis management.

The second source of tension between political and
military considerations is the issue of operational control:
tensions between the need for direct, positive, top-level
control of military operations in a crisis, and the need for
tactical flexibility and instantaneous decisionmaking at the
scene of the crisis. This tension arises from, and is a
symptom of, stratified interaction. The manner in which it
is handled by national leaders and the military chain of

command is a major determinant of whether or not stratified interactions become decoupled in a crisis.[19]

A fundamental and ubiquitous issue in civil-military relations is at what level in the chain of command should operational decisions on the employment of military forces and their weapons be made? Delegating decisionmaking authority to lower levels can deprive national leaders of the ability to control the momentum of a conflict and to coordinate military operations with diplomatic initiatives. On the other hand, centralizing decisionmaking authority in the hands of national leaders can rob on-scene forces of tactical flexibility, leaving them incapable of adapting to a rapidly changing situation and vulnerable to surprise attack. This is a dilemma inherent in the use of military force in crises. It cannot be resolved, it must be managed on an on-going basis every time a crisis occurs.

The nature of the interactions at the political and tactical levels can be quite different, and probably often are. Given a crisis in which national leaders on the two sides are pursuing strategies of coercive diplomacy, efforts by each side to enhance the credibility of its coercive threats or to counter the adversary's coercive threats by

---

[19]Although operational control tensions can also arise over the control of strategic level forces, this study will concentrate on tensions arising over the control of tactical level forces. On operational control tensions affecting strategic nuclear forces, see Sagan, pp. 99-139; Bracken, Command and Control of Nuclear Forces, pp. 196-202, 224-232.

deploying forces to the scene of the crisis will generate tactical level interaction. The commanders of those forces are constantly evaluating their tactical situation relative to adversary forces--assessing the seriousness of the threat they represent and the ability of his own forces to counter that threat. When the tactical situation is dynamic, with both sides acting to maintain or improve their tactical situation, the result is a test of capabilities at the tactical level. The test of capabilities is no less real for no shots having been fired. The tactical commanders are constantly calculating the likely outcome of an engagement with each new development in the tactical situation. Thus, even though political level interaction may be marked by coercive diplomacy, tactical level interaction can become a test of capabilities.

This illustrates why the level of control issue can be a source of tensions. The tactical situation can appear much different to the on-scene commander, operating under the guns of the adversary, than it does to top-level political leaders, negotiating with that same adversary. A military move by the adversary that is viewed as a political signal by national leaders can be viewed as an immediate military threat to the tactical commander. Under the conditions of a test of capabilities, a tactical commander is going to perceive an urgent need for as much decision-making authority as he can get from his chain of command.

At the same time, national leaders, particularly when
engaged in coercive diplomacy, are going to perceive an
urgent need for a high degree of direct control over the
actions of their military forces. The result is tension
between the tactical commander's need for flexibility and
initiative, and the politcal leader's need for close control
of military operations.

The third source of tension is wartime readiness:
tensions between performance of crisis missions and
maintaining or increasing readiness to perform wartime
missions. This is a tension between present operations and
possible future contingencies: the immediate political and
military objectives being pursued in a crisis conflicting
with the military objectives that would be pursued if the
crisis escalates to war. This tension arises for four
reasons, which may occur individually or together: dual
crisis-wartime tasking, replacement of crisis forces,
alliance commitments, and execution of contingency plans.

The first reason is that military forces do not
necessarily drop their wartime contingency tasking when
assigned to crisis operations. Their wartime tasking may
change to make it more compatible with their crisis tasking,
for example by assigning them wartime missions in the
vicinity of the crisis, but their wartime tasking is rarely
dropped altogether. The more capable and mobile a military
unit is, the more likely it is to retain significant wartime

contingency tasking while assigned to crisis operations. Forces equipped with dual-capable weapons--aircraft and missiles capable of carrying nuclear as well as conventional warheads--are most likely to have simultaneous crisis and wartime tasking.

Tensions arise from simultaneous crisis and wartime contingency tasking because the military requirements of the two missions can be significantly different. As an example, such differences in mission requirements would be particularly pronounced for U.S. forces when their crisis tasking entailed small-scale conventional operations against a much smaller nation (like Libya), while wartime contingency tasking entails large-scale conventional or nuclear operations against the Soviet Union.

The second reason why tensions arise between crisis objectives and wartime objectives is that forces detached from major operational units to respond to a crisis may be replaced by other forces in order to maintain readiness for wartime missions. For example, moving an aircraft carrier battle group out of the Western Pacific into the Arabian Sea for the possibility of operations against Iran can require that another carrier battle group be surged from its homeport in the United States to the Western Pacific in order to cover the wartime commitments of the first battle group. During an acute Soviet-American crisis, such surging of forces to replace crisis forces could send an inadvertent

signal of hostility to the Soviet Union--an example of the crisis security dilemma.

The third reason why tensions arise between crisis objectives and wartime objectives is that alliance commitments can inhibit forces deployed for support of allies in wartime from being used in a crisis. Alternatively, if forces normally committed to the defense of an ally are diverted to a crisis, other forces may have to be deployed from their home bases to avert sending an inadvertent signal of retrenchment to the ally. For example, U.S. forces in Korea are not readily available to respond to crises elsewhere in Asia.[20] Similarly, U.S. ground and air forces deployed in Western Europe to support NATO commitments are rarely employed for out-of-area contingencies. The frequent refusal of U.S. allies, notably NATO and Japan, to provide forces in support of U.S. military actions in crises--even crises affecting their interests more than American interests--further compounds crisis-wartime trade-offs for the United States.[21] In such

---

[20]Joseph F. Bouchard, "The American View of Korean Peninsula Security," in The Security Challenge in Northeast Asia: Report of a Conference (Stanford, CA: Stanford University, Northeast Asia-United States Forum on International Policy, December 1982), pp. 52-55.

[21]On NATO, see Gregory F. Treverton, "Global Threats and Trans-Atlantic Allies," International Security 5 (Fall 1980): 142-158; Edward A. Kolodziej, "Europe: The Partial Partner," International Security 5 (Winter 1980-1981): 104-131; Karl Kaiser, "NATO Strategy Toward the End of the Century," Naval War College Review 38 (January-February

situations the U.S. must divert forces from unilateral U.S. wartime contingencies in order to respond to a crisis without diverting forces from alliance wartime contingencies.

The fourth reason why tensions arise between crisis objectives and wartime objectives is that as a crisis escalates and military forces are placed at increasingly higher levels of readiness (DEFCON), initial preparations to execute wartime contingency plans commence. Certain of these preparations can be initiated by strategic level military commanders on the basis of authority delegated to them in mechanisms of delegated command, without an increase in DEFCON or other orders from political level leaders.[22] Increased surveillance of the potential enemy is an immediate measure. For conventional forces preparations for wartime operations include increased security measures to thwart enemy surveillance and deployment to wartime battle positions. If not carefully integrated with the diplomatic actions being taken to resolve the crisis, preparations to

---

1984): 69-82. On the NATO out-of-area issue, see A.W. DePort, "The North Atlantic Alliance: External Threats and Internal Stress," Naval War College Review 37 (November-December 1984): 71-79; William T. Tow, "NATO's Out-of-Region Challenges and Extended Containment," Orbis 29 (Winter 1985): 839-840. On Japan, see Joseph F. Bouchard and Douglas J. Hess, "The Japanese Navy and Sea-Lanes Defense," U.S. Naval Institute Proceedings 110 (March 1984): 88-97.

[22]Sagan, pp. 99-139; John Steinbruner, "An Assessment of Nuclear Crises," in Franklyn Griffiths and John C. Polanyi, eds., The Dangers of Nuclear War (Toronto: University of Toronto Press, 1979), pp. 34-49.

execute wartime contingency plans can touch off stratified escalation dynamics at the tactical and strategic levels of interactions.

In summary, tensions between political and military considerations are inherent in the use of military force as a political instrument. The first source of tension is conflict between political considerations and the needs of diplomatic bargaining, on the one hand, and military considerations and the needs of military operations, on the other. Tensions between political and military considerations arise because military forces must always be prepared for the possibility of combat even while carrying out political signaling missions. The second source of tension is the issue of operational control: tensions between the need for direct, positive, top-level control of military operations in a crisis, and the need for tactical flexibility and instantaneous decisionmaking at the scene of the crisis. This is a dilemma that cannot be resolved, it must be managed on an on-going basis every time a crisis occurs. The third source of tension is wartime readiness: tensions between performance of crisis missions and maintaining or increasing readiness to perform wartime missions. This tension arises for four reasons, which may occur individually or together: dual crisis-wartime tasking, replacement of crisis forces, alliance commitments, and execution of contingency plans.

Conclusion

This chapter began by describing the stratified interaction model and defining the theory of stratified interaction. The first corollary to the theory, decoupled interactions, was then presented. The theory of stratified interaction was then applied to crisis stability, producing the second corollary to the theory, that of stratified crisis stability. Extending this corollary to the problem of escalation resulted in the third corollary to the theory, stratified escalation dynamics. Finally, stratified interaction was used to explore the tensions that arise between political and military considerations when military force is used as a political instrument in crises.

With the theory of stratified interaction and its corollaries defined, we can now begin to explore the use of force as a political instrument in crises. The first task, which will be undertaken in the next chapter, is to examine the mechanisms of delegated command. These mechanisms are important in crisis management because the President, and even top-level military commanders, cannot possibly exercise real-time direct control over all the activities of the U.S. armed forces. The mechanisms of delegated command strongly influence the degree to which crisis interactions are stratified, the likelihood of stratified interactions becoming decoupled, and the intensity of the tensions between political and military considerations.

CHAPTER IV

MECHANISMS OF INDIRECT CONTROL

There are three major reasons for examining military command and control and the mechanisms of indirect control. First, and most important, there is always a danger that national leaders could lose control of events in a crisis, allowing crisis military operations to escalate uncontrollably to war.[1] The methods, capabilities, and limitations of military command and control are important factors in the ability of national leaders to maintain control over events. Second, the occurrence of stratified interaction in crises is largely a function of the military command and control procedures being employed to direct crisis military operations. Third, the nature and intensity

---

[1] See Alexander L. George, "The Development of Doctrine and Strategy," in Alexander L. George, David K. Hall, and William E. Simons, The Limits of Coercive Diplomacy (New York: Little, Brown and Co., 1971), pp. 8-15; John Steinbruner, "An Assessment of Nuclear Crises," in Franklin Griffiths and John C. Polanyi, eds., The Dangers of Nuclear War (Toronto: University of Toronto Press, 1980), p. 40; Phil Williams, Crisis Management (New York: John Wiley, 1976), pp. 94-134; Richard N. Lebow, Between Peace and War (Baltimore: Johns Hopkins University Press, 1981), p.287; Alexander L. George, "Crisis Management: The Interaction of Political and Military Considerations," Survival 26 (September/October 1984): 227-228.

of the tensions between political and military considerations that arise in crises are heavily influenced by military command and control procedures. Thus, greater familiarity with military command and control will enhance our understanding of escalation dangers in crises and the problems of crisis management.

The methods and problems of military command and control have received scant attention in studies of warfare and virtually no attention at all in studies of crisis management.[2] Virtually all the attention paid to command and control has been narrowly focused on technical issues-- maintaining reliable and rapid communications, improving information processing and display, and reducing vulnerability to enemy attacks and countermeasures. The recent spate of books and articles on the command and control of strategic nuclear forces has not corrected this deficiency. These studies have made an important

---

[2]For commentaries on this lack of attention, see Roger A. Beaumont, "Command Method: A Gap in Military Historiography," Naval War College Review 31 (Winter 1979): 61-74; Martin Van Creveld, Command in War (Cambridge, MA: Harvard University Press, 1985), p. 11. On the role of command and control systems in crisis management, see Davis B. Bobrow, "Communications, Command, and Control: The Nerves of Intervention," in Ellen P. Stern, ed., The Limits of Military Intervention (Beverly Hills: Sage, 1977), pp. 101-120; Phil Williams, "Crisis Management: The Role of Command, Control and Communications," RUSI Journal 128 (December 1983): 33-39; Garry D. Brewer and Paul Bracken, "Some Missing Pieces of the C^3I Puzzle," Journal of Conflict Resolution 28 (September 1984): 451-469. These articles address a neglected topic, but are focused on problems of maintaining connectivity.

contribution by exploring how operational problems in command and control systems can impair crisis management and escalation control.[3] Connectivity issues--ensuring that operational forces are reliably and securely linked with National Command Authority--are important, but there is much more to effective command and control.

The purposes of this chapter are to set military command and control in the context of theories on delegation and control in organizations, and to explain how delegation and control are exercised in the U.S. military command system. The first section will present organization and management theories on delegation and control, and show how they apply to military command and control. The second section will explore basic concepts of delegation and control used in the U.S. military command system. The third section will examine four of the mechanisms of indirect control: the alert system, standing orders, mission orders, and contingency plans. The final section will examine rules of engagement--the fifth mechanism of indirect control.

---

[3]Vice Admiral G.E. Miller, "Existing Systems of Command and Control," in Franklyn Griffiths and John C. Polanyi, eds., The Dangers of Nuclear War (Toronto: University of Toronto Press, 1979), pp. 50-66; Paul Bracken, The Command and Control of Nuclear Forces (New Haven, CT: Yale University Press, 1983); Scott D. Sagan, "Nuclear Alerts and Crisis Management," International Security 9 (Spring 1985): 99-139; Bruce G. Blair, Strategic Command and Control: Redefining the Nuclear Threat (Washington, DC: Brookings Institution, 1985); Ashton B. Carter, John D. Steinbruner, and Charles A. Zraket, eds., Managing Nuclear Operations (Washington, DC: Brookings Institution, 1987).

Delegation and Control in Organizations

The principle analytical concept currently employed for analysis of organizational behavior and the effect of that behavior on crisis interaction is the organizational process model.[4] That model has serious deficiencies when used as a tool for analysis of military command and control. Its weaknesses stem from two sources. First, it is based on a narrow and critically flawed conception of delegation and control in organizations. This weakness will be discussed in this section. Second, it fails to account for the manner in which military command and control is exercised, particularly the significant delegation of decisionmaking authority and the role of mechanisms of indirect control. This weakness will be addressed in the following section.

The organizational process model implicitly accepts the simple public administration distinction between policy-making and implementation. In the organizational process model, the President makes policy decisions and government organizations implement those those decisions. Organizational processes explain why the actions taken during implementation differ from the actions the President desired or expected when he made the decision. There is no provision in the model for government organizations to have

---

[4]Graham T. Allison, Essence of Decision: Explaining the Cuban Missile Crisis (Boston: Little, Brown and Co., 1971), pp. 78-96.

been delegated significant discretion in making operational decisions on how to implement policy decisions.[5]

This conception of implementation has long been discredited by political scientists and organization theorists. In their study of implementation, Jeffry L. Pressman and Aaron Wildavsky note that the distinction between policymaking and implementation is an oversimplification, and that "the passage of time wreaks havoc with efforts to maintain tidy distinctions" between the two functions. They contend that "In the midst of action the distinction between the initial conditions and the chain of causality begins to erode. Once a program is underway *implementers become responsible* both for the initial conditions and for the objectives toward which they are supposed to lead."[6] The implication of this is that the persons charged with carrying out a policy also have an

---

[5]Allison recognizes that "Government action requires decentralization of responsibility and power." However, his model is based on the observation that "The necessity for coordination and the centrality of foreign policy to the welfare of the nation guarantee the involvement of government leaders in the processes of the organizations that share power." This observation, while essentially correct, leads to an oversimplified model of governmental behavior as consisting of policy decisions by government leaders, and subsequent efforts by those leaders to control organizational routines so as to achieve desired results. The model thus ignores the deliberate delegation of decisionmaking authority and the role of such delegated authority in shaping policy. See Allison, pp. 85-87.

[6]Jeffrey L. Pressman and Aaron Wildavsky, Implementation (Berkeley: University of California Press, 1973), p. xxi.

important role in shaping the policy. Michael Lipsky had

made the same argument earlier in even stronger terms:

> There are many contexts in which the latitude of
> those charged with implementing policy is so
> substantial that studies of implementation should be
> turned on their heads. In these cases, policy is
> effectively "made" by the people who implement it.
> Where considerable discretion characterizes the jobs of
> people who implement public agency activities, people
> "make" policy in hidden concert with others in similar
> positions through their patterned responses to the
> situation and circumstances in which they find
> themselves.[7]

Thus, while organizational processes do have an influence on

policy outcomes, the organizational process model needs to

be revised to reflect the significant decisionmaking

authority--authority to define objectives and design

programs--delegated to certain officials in government

organizations.

The necessity for delegation of discretionary powers

and the coordination problems that can arise from this have

long been recognized in organization and management theory.

Chester Barnard recognized that delegation of discretion

results in policies being defined at all levels in an

organization, rather than just at the top.[8] According to

Anthony Downs, "At every level there is a certain

[7]Michael Lipsky, "Standing the Study of Policy
Implementation of Its Head," in W. Dean Burnham and Martha
W. Weinberg, eds., American Politics and Public Policy
(Cambridge, MA: MIT Press, 1968), p. 397.

[8]Chester I. Barnard, The Functions of the Executive
(Cambridge, MA: Harvard University Press, 1961), pp. 231-2.

discretionary gap between the orders an official receives from above and the orders he issues downward, and every official is forced to exercise discretion in interpreting his superior's orders."[9] This decentralization of decisionmaking is driven by limitations on the analytical capabilities of decisionmakers, which are rapidly exceeded as an organization increases in size and complexity. John W. Sutherland emphasizes this point: "Simply, as the scope of a decision-maker's authority increases (as the number of units for which he is responsible expands), the probability that he will make rational, accurate decisions about the properties of those programs decreases."[10] Thus, authority to define policies is diffused throughout organizations by the necessity of delegating discretionary powers in order to carry out top-level policy decisions.

In a useful refinement of this concept, Jay R. Galbraith has drawn a distinction between two methods of delegating decisions in organizations. In the first approach, rules, programs, and procedures are used to move repetitive decisions to lower levels in the organization without delegation of discretion. Decisionmaking by lower-level officials is guided by directives that specify the

_____

[9]Anthony Downs, <u>Inside Bureaucracy</u> (Boston: Little, Brown and Co., 1967), p. 134.

[10]John W. Sutherland, <u>Administrative Decision-Making: Extending the Bounds of Rationality</u> (New York: Van Nostrand Reinhold Co., 1977), p. 277.

actions to be taken those situations that can be anticipated in advance. According to Galbraith, "The primary effect is an information processing one--the elimination from hierarchical channels of communications concerning routine events. Rules serve the same function as habits for individuals. They preserve the scarce information processing, decisionmaking capacity for novel, consequential events."[11] Rule-governed delegation of decisionmaking is the type recognized in the organizational process model.

The second approach is to delegate discretionary decisionmaking authority. According to Galbraith, this is driven by an inability to anticipate situations for rule-governed decisions:

> The combination of rules and hierarchy, like hierarchy alone, is vulnerable to task uncertainty. As the organization's subtasks increase in uncertainty, fewer situations can be programmed in advance and more exceptions arise which must be referred upward in the hierarchy. As more exceptions are referred upward, the hierarchy will become overloaded. Serious delays will develop between the transmission of information upward and a response to that information downward. In this situation, the organization must develop new processes to supplement rules and hierarchy.
> As the task uncertainty increases, the volume of information from the points of action to points of decision making overload the hierarchy. In this situation, it becomes more efficient to bring the points of decision down to the points of action where the information exists. This can be accomplished by increasing the amount of discretion exercised by employees at lower levels of the organization.[12]

---

[11]Jay R. Galbraith, <u>Organization Design</u> (Reading, MA: Addison-Wesley, 1977), pp. 43-44.

[12]Ibid, p. 44. This is similar to Sutherland, p. 277.

Organizations typically use both methods of delegating decisions: rule-governed delegation of decisionmaking for standard, recurring situations, and discretionary delegation of decisionmaking for situations that cannot be anticipated.[13]

Discretionary delegation of decisionmaking raises the problem of ensuring that the decisions made by lower-level officials support the goals established by top-level officials. When this is not the case, delegated discretion results in "authority leakage," a divergence of goals between top-level and lower-level officials.[14] There thus arises an inherent tension between autonomy and control when discretionary delegation of decisionmaking is used to cope with uncertainty.

Various methods of control can be used by organizations. Galbraith emphasized two: professionalism, a reliance on professional training and socialization to ensure that officials make decisions that support organizational goals; and goal-setting, in which planned

---

[13]Ibid, p. 46. In practice, at least a small amount of discretion is allowed under rule-governed delegation--even if only discretion to determine which rules are applicable in specific situations--and at least minimal rules govern discretionary delegation. Thus, it is more accurate to portray delegation of decisionmaking as a spectrum ranging from highly rule-governed to highly discretionary, rather than as a dichotomy between the two types.

[14]Gordon Tullock, The Politics of Bureaucracy (Washington, DC: Public Affairs Press, 1965), pp. 142-93; Downs, pp. 134-5.

objectives are set and officials allowed to select
appropriate means for attaining the goals.[15] A scheme of
three categories of organizational controls is now commonly
used in organization and management studies: hierarchical
control, which include rules, procedures and directives
(which is Galbraith's rule-governed delegation); collegial
control, which is based on professional training and
identification; and nonhierarchical control, which is based
on internalization of the organization's norms and values.[16]
Organizations use all three of these categories to varying
degrees in order to maintain control under conditions of
discretionary delegation of decisionmaking.

These concepts of delegation and control have been
widely used in studies of school administration. As Kent
D. Peterson points out, the relationship between school
district officials and the principals of individual schools
highlights the issue of autonomy versus control:

> Functioning as the linkage between central office and
> classrooms as well as between parents and teachers,
> principals must keep resources, personnel, and students
> working efficiently toward organizational goals and

---

[15]Galbraith, pp. 45-6. Also see Arthur L. Stinchcombe, "Bureaucratic and Craft Administration of Production: A Comparative Study," Administrative Science Quarterly 4 (September 1959): 168-87.

[16]Tom K. Reeves and Joan Woodward, "The Study of Managerial Control," in Joan Woodward, ed., Industrial Organizations: Behavior and Control (London: Oxford University Press, 1970), pp. 37-56; William G. Ouchi, "A Conceptual Framework for the Design of Organizational Control Systems," Management Science 25 (September 1979): 833-848.

objectives. To do this, they must neither be so
tightly constrained that they cannot respond to
changing conditions, nor so loosely controlled that
they seek personal rather than organizational goals.
Superiors must afford the principal enough autonomy to
cope with unexpected problems or variable local
conditions, while still keeping schools in line. In
short, superiors seek an appropriate balance of control
and autonomy that will maximize organizational
effectiveness.[17]

Studies of school principals consistently find that they are

accorded significant autonomy and depict school districts

and the schools within them as being "loosely coupled."

Although all three forms of control are used in conjunction,

school principals typically are controlled largely by

collegial and nonhierarchical controls, rather than by

hierarchical controls.[18]

The organizational process model described by Graham

Allison in 1971 is incapable of accurately analyzing

organizational behavior in this type of situation. The

model would begin with the assumption that top-level school

district officials knew exactly what policies they wanted

implemented, ignoring the fact that they often do not know

what policies would be best for the specific conditions

[17]Kent D. Peterson, "Mechanisms of Administrative
Control over Managers in Educational Organizations,"
Administrative Science Quarterly 29 (December 1984): 573.

[18]Dan C. Lortie, "The Balance of Control and Autonomy
in Elementary School Teaching," in Amitai Etzioni, ed., The
Semi-Professions and Their Organizations (New York: Free
Press, 1969), pp. 1-53; Karl Weick, "Educational
Organizations as Loosely-Coupled Systems," Administrative
Science Quarterly 21 (March 1976): 1-19; Peterson, 573-97.

faced by each school and delegate substantial policymaking authority to the principals. The model would then assume that the sole function of school principles is to carry out school district policies, attributing policy differences among schools and policy disputes between principals and school district officials to organizational processes.

Allison's organizational process model leads to three serious analytical errors in this type of organizational setting. First, his model must assume that school district officials are always "right" and principals always "wrong" in order for the actions of principals to be attributable to organizational processes. The model needs to allow for the possibility that the policies decided upon by the implementers (principals) may be more appropriate or rational than those preferred by the policymakers (school district officials). Second, Allison's model must assume that the specific policy preferences expressed by school district officials when a dispute arises with a school principal were the policies that they had decided upon to begin with. The model needs to allow for the pos-sibility that school district officials may not have known exactly what policy they preferred until after they saw what the principal had decided upon. Third, Allison's model excludes the effect of variation in the environment of different decisionmakers, particularly the effect of differences in constituencies and political influences. School district

officials could well face much different political pressures than those faced by individual principals. The model needs to allow for such variations in the external environment. With these modifications, the organizational process model becomes applicable to an organization composed of several independent operating unit--such as a school district or operational military forces.

Studies of business management reveal patterns of delegation and control similar to those seen in public administration. Large business organizations face an inherent tension between the need to delegate decisionmaking and the need for centralized control, particularly when composed of diverse, autonomous operating units. They employ combinations of management controls (the business equivalent of the term organizational controls used in organization theory), including business variants of hierarchical, collegial, and nonhierarchical controls. The interesting point, however, is that businesses that decentralize decisionmaking authority typically perform better than do businesses that centralize decisionmaking. They are better able to respond to diverse and rapidly changing market, resource and regulatory conditions. Centralization of decisionmaking is widely regarded as stifling creativity, responsibility, loyalty, and entrepreneurship.[19]

[19]Kenneth A. Merchant, "The Control Function of Management," Sloan Management Review 23 (Summer 1982):

Several obsei ers have noted a trend toward
decentralization of decisionmaking in American businesses
that is increasing their productivity and competitive-
ness.[20] This includes allowing autonomous operating units
to define their own goals and strategies within broad,
flexible guidelines (that are often little more than a
corporate philosophy). Thus, in business management as well
as in public administra-tion there is recognition of the
need to balance delegation and control, and of the
advantages of decentralized decisionmaking authority

43-55; C.W.L. Hill and J.F. Pickering, "Divisonalization,
Decentralization and Performance of Large United Kingdom
Companies," Journal of Management Studies 23 (January 1986):
26-50; Michel Lebas and Jane Weigenstein, "Management
Control: The Roles of Rules, Markets and Culture," Journal
of Management Studies 23 (May 1986): 259-72; Vijay
Govindarajan, "Decentralization, Strategy, and Effectiveness
of Strategic Business Units in Multibusiness Organizations,"
Academy of Management Review 11 (October 1986): 844-856.
For business views, see Barrry A. Liebling, "Is it time to
(de)centralize?" Management Review 70 (September 1981): 14-
20; Don Collier, "Strategic Management in Diversified, Decen-
tralized Companies," Journal of Business Strategy 3 (Summer
1982): 85-89; J.H. Smith, "How Can Management Survive in a
Decentralized Environment?" Canadian Business Review,
Spring 1983, pp. 11-13; Robert E. Levinson, "Why Decentral-
ize?" Management Review 74 (October 1985): 50-53; Roy Hill,
"Centralization or autonomy: which way should a company
jump?" International Management, March 1986, pp. 17-18.

[20]In particular, see Thomas J. Peters and Robert H.
Waterman, Jr., In Search of Excellence (New York: Harper and
Row, 1982). Also see Michael Maccoby, The Leader (New York:
Simon and Schuster, 1981); Allan Cox, The Cox Report on the
American Corporation (New York: Delac rte Press, 1982);
Terrence E. Deal and Allan A. Kennedy, Corporate Cultures
(Reading, MA: Addison-Wesley, 1982); Rosabeth M. Kanter, The
Change Masters (New York: Simon and Schuster, 1983).

when used with appropriate--primarily collegial and nonhierarchical--control mechanisms.[21]

In summary, the study of organizational behavior must account for significant delegation of decisionmaking authority. Delegation of decisionmaking is driven by the limits on decisionmaking, which cause decision-making by top-level officials to deteriorate as the size and complexity of the organization increase. Delegation of decisionmaking can range from highly rule-governed, for standard, repetitive situations, to highly discretionary, for situations that cannot be anticipated. Three types of control mechanisms are used in various combinations: hierarchical (rules and procedures), collegial (professionalism), and nonhierarchical (organizational and societal norms and culture). Tension between autonomy and control is always present, particularly in organizations consisting of numerous independent operating units. Studies in public administration and business management repeatedly show that optimum results are achieved with decentralized decisionmaking combined with appropriate controls.

---

[21]The trend toward decentralization in American business management has been accompanied by scathing criticism of the highly centralized management style that Robert S. McNamara adopted at Ford and brought with him to the Department of Defense in 1961. See Robert E. Levinson, "The high cost of remote control management," Management Review 72 (April 1983): 12-20; and Gordon Pearson, "Business strategy should not be bureaucratic," Accountancy, April 1986, pp. 109-12.

Military Command and Control

Command and control is generally viewed in narrow terms of organizational structures, communications systems, and information processing systems. Attention to communications and information processing is certainly warranted, for the effective exercise of command and control is crucially dependent on the commander's ability to build a clear picture of the operational situation. As Martin van Creveld notes: "From Plato to NATO, the history of command in war consists essentially of an endless quest for certainty—certainty about the state and intentions of the enemy's forces; certainty about the manifold factors that together constitute the environment in which the war is fought, from the weather and the terrain to radioactivity and the presence of chemical warfare agents; and, last but definitely not least, certainty about the state, intentions and activities of one's own forces."[22] This imperative applies to crisis miltiary operations as well as to wartime operations. Even before the shooting starts, it is crucial that the chain of command up to top-level national leaders have as clear a picture as possible of the situation at the scene of a crisis.

The attention paid to communications and information processing systems has overshadowed and distracted attention

---

[22] Creveld, p. 264.

from the fundamental command and control functions supported
by those systems. From a crisis management perspective,
ensuring that the radios and computers operate properly is
the lesser problem--knowing how to effectively control
military operations with them is the more difficult
problem. Outside of the military training courses that
train officers for leadership and command, little attention
is paid to the methods, procedures, and mechanisms of
command. Rapid advances in communications and information
processing technology are having a tremendous impact on
military command and control, but the manner in which the
systems are employed operationally is still primarily a
function of the command and control philosophy held by the
personnel using the systems.

Prior to examining command and control procedures and
mechanisms, it is necessary to understand the fundamental
concepts of command and control as they are defined by the
military. The Joint Chiefs of Staff define command and
control as "the exercise of authority and direction by a
properly designated Commander over assigned forces in the
accomplishment of his mission." The essence of command is
authority and responsibility. Authority is the power to
direct the operations and movements of the forces under
one's command. Responsibility is being held accountable for
the performance and well-being of the forces and men under
one's command. A key tenant of military leadership is that

while authority can be delegated, responsibility cannot.[23] That is, a commander can delegate authority over a portion of his forces to a subordinate, but retains responsibility for those forces.

A closely related principle--unity of command--states that if a commander is given responsibility for forces, he must have authority over them commensurate with that responsibility. This is the principle of command that is most difficult to uphold. Military commanders and civilian authorities alike face a constant temptation to restrict the authority of subordinates even while holding them account-able for the actions of their forces.[24] Ambiguous delegation of authority almost invariably leads to diffusion of responsibility, a phenomenon highly visible when military operations fail. Diffusion of authority and responsibility can be unintended but nonetheless deleterious side-effects of modern communications systems.

---

[23]See Joint Chiefs of Staff, Department of Defense Directory of Military and Associated Terms, JSC Publication No. 1 (Washington, DC: U.S. Government Printing Office, 1979); United States Navy Regulations, 1973 (Washington, DC: U.S. Government Printing Office, 1973), p. 13; Admiral James D. Watkins, "The Principle of Command," U.S. Naval Institute Proceedings 109 (January 1983): 32-33.

[24]See Admiral Harry D. Train III, "Decision Making and Managing Ambiguity in Politico-Military Crisis," in James G. March and Roger Weissinger-Baylon, eds., Ambiguity and Command: Organizational Perspectives on Military Decision Making (Marshfield, MA: Pitman Publishing, 1986), p. 307; Lieutenant D.J. Marchall, "Communications and Command Prerogative," U.S. Naval Institute Proceedings 100 (January 1974): 29-33.

Another important principle of military command is the distinction is drawn between command and control. The two functions can be exercised separately. A commander can delegate control over forces he commands. He retains responsibility for the forces, but grants the subordinate authority over them. The subordinate commander is then responsible for the forces under his control. For example, when a company commander sends a squad out on patrol, he delegates control of the men to the squad leader, but is still in command of them. When military command functions in accordance with this principle, the superior commander exercises authority over the subordinate commander, not over the forces placed under the control of the subordinate commander. The military chain of command is founded on the principle of delegating control while retaining command.

A key point that is often missed in studies of command and control is that this distinction between command and control starts with the commander in chief of the armed forces--the President. Under the United States constitution, the President is, in principle, in command of every unit and individual member of the U.S. armed forces at all times. However, while retaining that command, he delegates substantial control over the armed forces to the Secretary of Defense, who in turn delegates substantial control to subordinate commanders. Presidential delegation of control over military forces varies widely in extent and

method. For example, the President retains tight control over release of nuclear weapons, but exercises very little control over routine peacetime military operations. Understanding the distinction between command and control and the delegation of control to military commanders is thus essential for understanding how crisis military operations are controlled.

This section will examine four aspects of military command and control. First, the tension between delegation and control in the military command system will be explored, presenting the arguments for and against centralized control of military operations. Second, the methods of exercising control--direct versus delegated, and positive versus by negation--will be explained. Third, the differences between initiatory actions and contingent responses will be defined. Finally, the authority of U.S. Navy commanding officers will be discussed and illustrated with historical examples.

Delegation and Control

Tension between delegation of discretionary decisionmaking authority and maintenance of control is always present in organizations, particularly in large organizations consisting of numerous independent operating units. There is no better example of such an organization than the United States armed forces. The Department of

Defense is by far the single largest organization in the
U.S. government, and it consists of innumerable independent
operating units with varying degrees of autonomy. Moreover,
unlike any other organization in the U.S., the armed forces
are charged with a mission crucial to the survival of the
nation. Organization theory would thus lead to the
expectation that there would be significant tension between
delegation of discretionary decisionmaking authority and
maintenance of control in the military chain of command.

That expectation is borne out by substantial empirical
evidence. In fact, centralization versus decentralization
in the control of military operations had been a major issue
in American civil-military relations and the design of the
U.S. military command system since the National Security Act
of 1947. The Truman-MacArthur dispute during the Korean War
and military dissatisfaction with Johnson Administration
"micro-management" of the air war against North Vietnam are
only two of the most prominent examples of such tensions.
It is thus important to understand the roots of such
tensions--the reasons advanced for centralized control and
the opposing reasons advanced for decentralized delegation
of control.

Three primary reasons or explanations for the trend
toward centralized control of military operations have been
advanced. The first is that the increasing complexity of
warfare and concomitant specialization of military forces

has increased the need for centralized control over military

operations in order to effectively coordinate diverse

weapons and units.[25] This affects almost every aspect of

military operations and is widely recognized within the

military. Since World War II U.S. Air Force doctrine has

called for centralized (theater level) control of tactical

air power.[26] Similar control problems arise in amphibious

operations and combined arms ground operations. The second

reason for centralized control, the one most important to

civilian leaders, is the threat of nuclear war.[27] This, in

turn, leads to the emphasis on top-level control of military

operations for escalation control and crisis management.[28]

---

[25]Sir Solly Zuckerman, "Judgement and Control in Modern Warfare," Foreign Affairs 40 (January 1962): 203-5; Creveld, pp. 236-7.

[26]For a description of Air Force command doctrine for tactical air power, see General William Momyer, Airpower in Three Wars (Washington, DC: U.S. Government Printing Office, 1978).

[27]Rear Admiral Donald T. Poe, "Command and Control: Changeless--Yet Changing," U.S. Naval Institute Proceedings 100 (October 1974): 23-4; Roger A. Beaumont, "The Paradoxes of C^3," in James H. Buck and Lawrence J. Korb, eds., Military Leadership (Beverly Hills: Sage, 1981), p. 123-4; Barry R. Posen, "Inadvertent Nuclear War? Esclation and NATO's Northern Flank," International Security 7 (Summer 1982): 28-54; Steinbruner, p. 40; Williams, pp. 94-134; Lebow, p.287.

[28]See George, "Development of Doctrine and Strategy," pp. 8-15, and "Crisis Management," pp. 227-228; Richard G. Head, Frisco W. Short, and Robert C. McFarlane, Crisis Resolution: Presidential Decision Making in the Mayaguez and Korean Confrontations (Boulder, CO: Westview Press, 1978), pp. 246-51.

The third, and most controversial, reason for centralization is the "force multiplier" concept. The contention is that effective command and control systems in effect multiply the combat utility of available forces by allowing them to be rapidly applied where they are most needed or where they can achieve the greatest results.[29] In principle, the higher the level in the chain of command at which control is centralized, the broader the force multiplier effect can be applied. Of these three reasons for centralized control, escalation control and crisis management concerns are the most important factors prompting civilian control of military operations.

Two primary reasons are advanced for decentralized control of military operations. First, the ability of top-level decisionmakers to effectively exercise close control of military operations is severely constrained by limits on decisionmaking and information processing. Top-level decisionmakers can be overwhelmed by information overload, preventing effective assessment of tactical options. They may not have sufficient time to effectively control multiple operations, or may have their attention diverted by one aspect of the operations, neglecting others. They almost invariably do not understand the complexities of modern warfare, which can make even a small-scale operation

---

[29]For a discussion of the force multiplier concept, see Beaumont, "The Paradoxes of C^3," pp. 116-20.

impossible to effectively control from the White House. As Ernest R. May has pointed out, there have been quantum leaps in the level of knowledge that the President must have of military forces in order to be able to employ them effectively. Communications channels typically become overloaded with the vast amounts of information needed to exercise close control of military operations, causing excessive delays in decisionmaking and transmission of orders to operating forces. Compounding these problems, the quality of modern communications systems can give top-level officials a false sense of having complete information and being in control.[30] That these problems should arise in centralized control of military operations is not surprising because, as was noted above, essentially the same phenomena drive decentralization in all types of large organizations.

The second set of arguments for decentralized control of military operations are based on the on-scene commander's superior ability to control the employment of his forces. His information about the current tactical situation is normally superior that of his superiors. The on-scene commander requires initiative and flexibility to

---

[30]Ernest R. May, "Eisenhower and After," in Ernest R. May, ed., The Ultimate Decision: The President as Commander in Chief (New York: George Braziller, 1960), pp. 233-5; Captain W.T.T. Pakenham, "The Command and Control of Naval Operations: Principles and Organisation," Naval Forces 7 (1/1986): 50; Beaumont, "Command Method," pp. 65-7, and "The Paradoxes of C³," p. 133; Creveld, p. 247-51; Poe, pp. 28-9.

effectively cope with the "fleeting opportunities and sudden dangers" of combat, to use Edward N. Luttwak's apt expression. In recognition of the importance of initiative and flexibility, the German army has since the eighteenth century based its tactical doctrine and command procedures on the concept of auftragstaktik, which emphasizes granting subordinate commanders the maximum possible freedom of action in carrying out assigned missions. Only the on-scene commander can fully appreciate and adapt to the inevitable "friction" in military operations, the multitude of problems that shape the execution of military plans. Centralized control of military operations can stifle initiative, weaken morale, erode authority, and cause diffusion of responsibility.[31] These are the concerns in the minds of military commanders when the White House gets on the radio to dictate their tactics.

---

[31]Colonel H.A. Hadd, "Orders Firm But Flexible," U.S. Naval Institute Proceedings 88 (October 1962): 87-8; Admiral Thomas B. Hayward, "An Ex-CNO's Reflection of the Garbage Can Theory of Naval Decision Making," in James G. March and Roger Weissinger-Baylon, eds., Ambiguity and Command: Organizational Perspectives on Military Decision Making (Marshfield, MA: Pitman Publishing, 1986), p. 267; Edward N. Luttwak, Strategy: The Logic of War and Peace (Cambridge, MA: Belknap Press, 1987), p. 13; Dan Horowitz, "Flexible Responsiveness and Military Strategy: The Case of the Israeli Army," Policy Sciences 1 (Summer 1970): 191-205; Beaumont, "The Paradoxes of C³," pp. 123, 133; Creveld, p. 269-70; Poe, p. 28-9; Train, p. 302-3; Blair, Strategic Command and Control, p. 75. For a description of auftragstaktik, see Lieutenant Colonel Walter von Lossow, "Mission-Type Tactics versus Order-Type Tactics," Military Review 57 (June 1977): 87-91.

The argument is frequently made that improved communications and information processing systems can overcome most of the problems that constrain top-level control of military operations. Such optimism is not supported by historical evidence. The historical trend has been for increases in the scale, speed, and complexity of warfare to exceed the ability of command and control systems to keep higher level commanders fully in control.[32] Martin van Creveld has reached the same conclusion:

> Taken as a whole, present-day military forces, for all the imposing array of electronic gadgetry at their disposal, give no evidence whatsoever of being one whit more capable of dealing with the information needed for the command process than were their predecessors a century or even a millennium ago. Though modern technical means undoubtedly enable present-day command systems to transmit and process more information faster than ever before, regardless of distance, movement and weather, their ability to approach certainty has not improved to any marked extent. Nor, given the fact that this goal has proved elusive through every one of the many revolutions in organization, technology and procedure that have taken place in the past, does there appear to be much hope of achieving it in the foreseeable future.[33]

---

[32]Beaumont, "Command Method," pp. 62-5; Lieutenant Colonel Kenneth F.A. Openchowski, "The Role and Location of the Commander: How Will They Be Affected by C^3 Facilities Available in the 1980s?" Military Review 57 (April 1977): 12-19.

[33]Creveld, pp. 265-6.

The trend toward complexity is particularly acute in naval operations, which are conducted with a wider array of sensors, platforms, and weapons than any other type of military operation--making it the most difficult form of warfare to explain to civilian leaders and advisors.[34] There is thus little reason to expect that innovations in communications and information processing systems will sovle crisis command and control problems.

Methods of Control

The military chain of command, from the President down to the lowest levels, is founded upon the principle of delegating control of forces to subordinate commanders. The methods of exercising control cover a "tightness of control" spectrum ranging from positive direct control at the tight end to autonomous delegated control at the loose end.[35]

---

[34]See Karl Lautenschlager, "Technology and the Evolution of Naval Warfare," International Security 8 (Fall 1983): 3-51; Admiral James D. Watkins, "The Maritime Strategy," U.S. Naval Institute Proceedings 112 (January 1986), special supplement, The Maritime Strategy, pp. 12-14. Modern naval command and control systems add to the complexity of warfare, rather than simplifying its control, by allowing a wider variety of weapons covering larger ocean areas to be brought to bear on the enemy. This has lead to enemy command and control systems becoming a primary wartime target. See Norman Friedman, "C^3 War at Sea," U.S. Naval Institute Proceedings 103 (May 1977): 126-41.

[35]The terms used in this section are derived from terms used by the three services, but are not the exact terms used by any of the services. There are two reasons for this. First, terminology varies widely among the three services

The tightest form of control is positive direct control. In this method, communications links with operational forces are used to control their movements and actions on a real-time basis. Positive direct control allows subordinate commanders the least amount of initiative and flexibility: movements and actions are taken only on direct orders. If the on-scene commander wants to take an action other than that ordered by higher authority, he must request and receive permission to do so before initiating the action. The effectiveness of positive direct control is crucially dependent on communications connectivity and having the full, undivided attention of higher authority. This form of control is rarely used for wartime or large-scale crisis operations because it is cumbersome and incapable of keeping pace with a rapidly changing tactical situation.

Toward the center of the tightness of control spectrum is a method that will be referred to as direct control by negation. As in positive direct control, communications links with operational forces are used to control their movements and actions on a real-time basis. However, the on-scene commander is delegated partial authority to control his forces. The scope of the on-scene commander's decisionmaking authority is defined in his mission orders,

---

and the joint commands despite efforts by the JCS to standardize it. Second, the terminology used by the services is much more complex than the scheme used here, employing myriad terms for different types of control.

and can vary widely. The key feature of this method of control is that the on-scene commander reports his proposed course of action to higher authority, and then carries it out unless it is specifically vetoed. Like positive direct control, the effectiveness of direct control by negation is crucially dependent on communications connectivity and having the full, undivided attention of higher authority.

Higher authorities can specify either or two forms of direct control by negation: tight or loose. In the tight form, the on-scene commander reports proposed actions before initiating them (except when immediate action is needed in an emergency). In the loose form, the on-scene commander initiates action before reporting it. The loose form of direct control by negation is the method of control that military commanders typically prefer when they must be placed under direct control.

Toward the loose end of the tightness of control spectrum are the various forms of delegated control. In this method of control the immediate commander of a force is delegated direct control over its operations. The commander is issued orders to perform a certain mission and then allowed to carry out that mission on his own initiative. The scope of his authority and the actions he is permitted to take in pursuit of the mission are spelled out in his mission orders, and can range from granting him wide freedom of action to restricting him to a specific plan of action.

When direct communications links are available, monitored delegated control is the form commonly employed. In monitored delegated control, the on-scene commander is required to keep his superiors informed of the status of his forces, the progress of his mission, and his operational intentions through periodic situation reports and, if possible, real-time reports of crucial information. The chain of command intervenes in the conduct of the operation only when absolutely necessary to ensure that it supports the overall strategy being pursued or to correct serious (mission-threatening) errors by the on-scene commander. Control by negation is often employed in monitored delegated control, as well as under direct control by negation. The difference between the two methods is that monitored delegated control allows the on-scene commander greater freedom of action than does direct control by negation. Monitored delegated control is viewed by most military officers as the ideal method of control, striking an effective balance between autonomy and control.

When direct communications links are not available or not feasible, autonomous delegated control is the form of control that is employed. In autonomous delegated control the on-scene commander is given his mission orders and is not expected to report again to higher authority until he successfully completes or aborts the mission. This type of control is necessary in covert operations, such as by

special forces, when stealth is crucial to the effectiveness
of a platform, such as attack submarines, and in situations
where electronic emissions must be curtailed, such as in a
surprise attack or when deception is used in battle. This
method of control is not often favored by military
commanders because it lacks flexibility, and is therefore
only used when absolutely necessary.

The approach the United States armed forces have taken
is to rely on a flexible combination of direct and delegated
methods of control. The balance between autonomy and
control is based on the nature of the operations being
conducted and the tactical environment--including the
political environment. In some cases, different methods of
control can be used in conjunction. For example, an on-
scene commander could be under monitored delegated control,
but have certain tactical options placed under positive
direct control. Admiral Joseph Metcalf, commander of the
Grenada invasion force in 1983, used this approach to allow
his subordinate commanders maximum freedom of action while
retaining control over weapons with the greatest destructive
power (such as attack aircraft and naval gunfire).[36] Forces

---

[36]Vice Admiral Joseph Metcalf III, "Decision Making in
the Grenada Rescue Operation," in James G. March and Roger
Weissinger-Baylon, eds., Ambiguity and Command:
Organizational Perspectives on Military Decision Making
(Marshfield, MA: Pitman Publishing, 1986), p. 281-2. Also
see Poe, p. 28, on flexible command procedures.

can be rapidly shifted between the different methods of control as the tactical situation dictates.

Initiatory Actions and Contingent Responses

Another concept useful for understanding military command and control is the distinction between initiatory actions and contingent responses.[37] Initiatory actions are taken to initiate a new course of action intended to achieve specific objectives. They may be directed by higher authority (directly or in advance), or taken by the on-scene commander on his own authority (based on general operational guidance and tactical doctrine). Although initiatory actions usually are taken on the basis of an assessment of the overall political-military environment and the local tactical situation, they are not the product of rules designating a particular response to a specific action. The on-scene commander is taking the initiative, acting rather than reacting.

Contingent responses, on the other hand, are actions taken in response to specific actions. The taking of such actions is contingent upon the prior occurrence of specified actions or the existence of specified tactical

---

[37]As in the previous section, these terms are derived from terms used by the three services, but are not the exact terms used by any of the services. The basic idea conveyed by these terms is deeply ingrained in military thought on command, and underlies certain commonly-used procedures (such as rules of engagement).

circumstances. Contingent responses are rarely ever fully automatic, they usually require a deliberate operational decision by the on-scene commander.[38] The principle variable is the level in the chain of command at which various contingent responses can be ordered. Decisions on the use of force governed by rules of engagement are a form of contingent response. Other contingent responses, usually broader in scope, can be included in operations orders, operations plans, and contingency plans. As a general rule, the broader the scope and the greater the level of violence involved in a contingent response, the higher up the chain of command the decision to employ that response must be made.

This distinction between initiatory actions and contingent responses cannot be pushed too far--they can be

---

[38]Certain naval weapon systems have automatic reaction capabilities: they can proceed from initial detection of a target to firing of weapons without an orders from an operator. The best-known example is the MK 16 Close-In Weapon System (the white-domed 20mm Vulcan gun installed on almost all U.S. Navy warships), which in the automatic mode will detect, track, and fire on any air target that meets its engagement parameters. The new Aegis combat system also has sophisticated automatic reaction and engagement capabilities. See Captain Joseph L. McClane, Jr. and Commander James L. McClane, "The Ticonderoga Story: Aegis Works," U.S. Naval Institute Proceedings 111 (May 1985): 118-29; Thomas B. Blann, "The State of Surface Antiair Warfare," U.S. Naval Institute Proceedings 111 (November 1985): 133-37. However, all such systems have semi-automatic modes that require an operator to make the decision to engage a target, and ships are usually given detailed guidance on the operation and programming of automatic systems so as to ensure that they are employed in accordance with the rules of engagement.

difficult to distinguish in a fast-paced tactical environment. For example, a retaliatory attack can be either a contingent response authorized in operational guidance (other than the rules of engagement), or an initiatory action ordered by national leaders--even though taken as a reprisal for a specific hostile act. Contingent responses and initiatory actions can have the same political-military intentions and objectives. A retaliatory attack, to continue the example, could have as its purpose the signaling of a coercive threat regardless of whether it was a contingent response or an initiatory action. The primary difference between these two types of military actions is that operational commanders can, on their own authority, execute contingent responses under specified tactical circumstances, whereas appropriate higher authority (normally the President in peacetime) must approve initiatory actions.

An important feature of the flexible system of direct and delegated command used by the U.S. military is that authority to order contingent responses can be delegated separately from authority to order initiatory actions. Orders for a particular mission can specify positive direct control of initiatory actions, while at the same time employing monitored delegated control of contingent responses. This approach is particularly useful when ships are deployed to a tense crisis situation marked by a high-threat tactical environment: top-level decisionmakers

retain control of actions most likely to be escalatory

(initiatory actions) while providing on-scene decision-

makers the tactical flexibility they need to defend

themselves (contingent responses). As always, a careful

balance must be struck between direct and delegated command

when using this differentiated approach. The important

point is that the decision on direct versus delegated

command is not an "all or nothing" proposition--flexible

combinations of the two methods are possible.

Authority of Commanding Officers

An important aspect of United States Navy command and

control philosophy, which has no equivalent in the other

services, is the extremely high delegation of authority

granted to the commanding officers of ships and other

operational commanders. This is a tradition in the U.S.

Navy, one inherited from the Royal Navy and centuries-old

traditions of the sea. During the age of sail, when it

could take longer to exchange letters with the homeland than

to fight a small war, British and American naval officers

often played important foreign policy roles on their own

initiative. British naval officers were under standing

orders to "act in the best interests of the Queen" when

deployed to distant stations, granting them freedom of

action to handle situations not covered in their sailing

orders. During the nineteenth century, naval officers

played an important role in U.S. foreign Policy, often with considerable autonomy.[39] This tradition exerts a strong influence on Navy command and control philosophy today.

Because the actions of individual commanding officers could have a major impact on the management of a crisis, it is important to understand their authority as spelled out in United States Navy Regulations. Both the 1948 and 1973 editions state the following:

> The responsibility of the Commanding Officer for his command is absolute, except when, and to the extent, relieved therefrom by competent authority, or as provided otherwise in these regulations. The authority of the Commanding Officer is commensurate with his responsibility, subject to the limits proscribed by law and these regulations.[40]

A commanding officer's authority and responsibility are thus "absolute," limited only by law and Navy Regulations.

---

[39] See Samuel Eliot Morison, "Old Bruin": Commodore Matthew C. Perry, 1794-1858 (Boston: Little, Brown and Co., 1967); Kenneth J. Hagan, American Gunboat Diplomacy and the Old Navy, 1877-1889 (Westport, CT: Greenwood Press, 1973); Robert E. Johnson, Far China Station: The U.S. Navy in Asian Waters, 1800-1898 (Annapolis, MD: Naval Institute Press, 1979); William N. Still, Jr., American Sea Power in the Old World: The United States Navy in European and Near Eastern Waters, 1865-1917 (Westport, CT: Greenwood Press, 1980); Frederick C. Drake, The Empire of the Seas: A Biography of Rear Admiral Robert Wilson Shufeldt, USN (Honolulu: University of Hawaii Press, 1984); David F. Long, Gold Braid and Foreign Relations: Diplomatic Activities of U.S. Naval Officers, 1798-1883 (Annapolis, MD: Naval Institute Press, 1988).

[40] U.S. Department of the Navy, United States Navy Regulations, 1948 (Washington, DC: U.S. Government Printing Office, 1948), p. 75; U.S. Department of the Navy, United States Navy Regulations, 1973 (Washington, DC: U.S. Government Printing Office, 1973), p. 13.

Commanding officers are, of course, required to carry out lawful orders from superiors, but <u>Navy Regulations</u> even has a provision covering situations in which a commanding officer believes he must act contrary to his orders:

> A Commanding Officer who departs from his orders or instructions, or takes official action which is not in accordance with such orders or instructions, does so upon his own responsibility and shall report immediately the circumstances to the officer from whom the prior orders or instructions were received.[41]

This is a key provision: It permits a commanding officer, under unanticipated extraordinary circumstances, to exercise initiative--even when contrary to his orders--so long as he immediately informs his chain of command of his action. In practice, commanding officers are extremely cautious about taking such an action. Their professional training and experience instill in them strong respect for the chain of command and orders issued by higher authority.

Two examples of Navy officers acting on their own initiative, in one case contrary to orders, will illustrate how the provisions of <u>Navy Regulations</u> are applied in practice. The first case occurred in July 1953, two days after the end of the Korean War. On July 29, 1953, a U.S. Air Force RB-50 reconnaissance plane patrolling in international airspace over the Sea of Japan was shot down by Soviet fighters about thirty miles off the coast of the Soviet Union (apparently in retaliation for the downing of a

---

[41]Ibid.

Soviet transport over North Korea two days earlier). All but one of the crew were able to bail out, and several were spotted in the water by American search planes.[42] Six ships, a cruiser and five destroyers under the command of Vice Admiral Walter G. Schindler, were detached to rescue the survivors. Vice Admiral Schindler's orders did not specify how close to the Soviet coast he was allowed to search, so he sent the following message to his superiors:

> Request you relay via appropriate channels to Russian authorities that if warranted by situation I intend to take my ships as close to Russian territory as is necessary to recover the airmen from the crashed aircraft and that furthermore, in the event I do, I will brook no interference.[43]

By the time Vice Admiral Schindler received a response directing him to remain clear of Soviet territorial waters, he had already recovered the only survivor that could be located. As it turned out, the lone survivor was found in international waters, none of Vice Admiral Schindler's ships entered Soviet territorial waters, and there was no harassment of his force by Soviet ships or planes.[44]

---

[42]James A. Field, Jr., History of United States Naval Operations: Korea (Washington, DC: U.S. Government Printing Office, 1962), pp. 457-9.

[43]Rear Admiral Daniel V. Gallery, The Pueblo Incident (Garden City, NY: Doubleday and Co., 1970), p. 57; Vice Admiral Walter G. Schindler, letter to author, March 19, 1988. Admiral Schindler confirmed that he sent the message reproduced in Gallery's book.

[44]Field, p. 459; Gallery, p. 58; Schindler, letter to author, March 19, 1988.

This episode illustrates two points. First, it illustrates monitored delegated control in practice. Admiral Schindler informed his superiors of his intentions, allowing control by negation. Second, it illustrates the authority of a Navy commander. Vice Admiral Schindler had the authority to order the actions taken by his force, used a message to his superiors stating his intentions in order to clarify ambiguous orders, and, by keeping his force clear of Soviet territorial waters, acted with appropriate caution in the absence of a timely response.

The second case occurred in Zanzibar, an island nation off the eastern coast of Africa (now part of Tanzania), in January 1964. Zanzibar, a former British protectorate, had gained independence on December 10, 1963. On January 12, 1964, African rebels overthrew the government of sultan Seyyid Jamshid Bin Abdullah in a bloody coup. Little was known about the rebel group or its intentions, leading to grave concern for the safety of foreigers on the island.[45]

The United States immediately ordered the destroyer USS Manley (DD 940) to proceed to Zanzibar from Mombassa, Kenya, where it had been making a port visit, and to establish a visible presence off the port city of Zanzibar, the island's

---

[45]"The Sultan Banished," Tanganyika Standard (Dar Es Salaam), January 14, 1964, p. 1; "African Revolt Overturns Arab Regime in Zanzibar," New York Times, January 13, 1964, p. 1; Jules Davids, The United States in World Affairs, 1964 (New York: Harper and Row, 1965), p. 278.

capital. While the destroyer was en route, however, a rebel leader broadcast a warning that they would not tolerate interference by foreign powers. In response to this development, U.S. authorities in Washington (the Secretary of Defense was directly involved) first ordered Manley to remain out of sight over the horizon from the island, then later cancelled Manley's orders entirely and recalled the destroyer.[46]

Manley received the first message, ordering her to remain over the horizon, just as she arrived at Zanzibar and established radio communications with the American embassy on the island. The sixty-three Americans on the island, including the staff of a NASA space tracking station, had fled to the English Club on the Zanzibar city waterfront. The American Charge d'Affairs, Frederick P. Picard, informed the destroyer that their situation on the island was desperate and that they were in grave danger. He requested that Manley evacuate the Americans immediately. Manley's Commanding Officer, Commander Robert Ruxton, reported the evacuation request to his immediate superior, Rear Admiral

---

[46]Ibid.; "Coup in Zanzibar Stirs U.S. Concern," New York Times, January 14, 1964, p. 1; Captain John H. Carmichael, Assistant Director of the Fleet Operations Division, Office of the Chief of Naval Operations in 1964, letter to author, March 8, 1988; Captain Joseph E. Murray, Jr., Executive Officer of USS Manley (DD 940) in 1964, letter to author, 31 March 1988; Captain Murray, letter to his wife, January 14, 1964 (Provided to author by Captain Murray).

Arnold F. Schade, Commander Middle East Force. Rear Admiral
Schade, acting on his own authority and contrary to the
orders, that had been sent from Washington, ordered <u>Manley</u>
to evacuate the Americans on the island.[47]

<u>Manley</u> sent the ship's Executive Officer, Lieutenant
Commander Joseph E. Murray, Jr., ashore--unarmed and in a
white uniform--to organize the evacuation. Murray and
Picard negotiated with the rebel leaders, assuring them that
the U.S. warship would not interfere in Zanzibar's internal
affairs. At one point in the talks, the rebel leader, John
Okello, put a gun to Murray's head and threatened to kill
him if the U.S. ship did not depart. Murray and Picard
persevered, gaining permission from the rebel leaders to
evacuate American women and children. When they returned to
the English Club, Murray and Picard convinced the rebel
guards that they had permission to evacuate all of the
Americans. Murray succeeded in ferrying all of the
Americans (and several non-U.S. citizens) to Manley before
the rebel leaders discovered what had happened. <u>Manley</u>
embarked 61 Americans and 30 other foreign nationals, and
departed for Dar Es Salaam, Tanganyika (leaving behind
Picard and the third secretary to handle relations with the
new government). Washington was unable to participate
because it did not have radio or cable communications with

---

[47]Ibid.

Zanzibar, and did not learn of the evacuation until after

Manley departed. After Manley departed Zanzibar, the order

from Washington cancelling the mission was received.[48]

This episode clearly illustrates the authority of Navy

commanding officers, including their authority to disregard

orders from superiors when the situation warrants. Rear

Admiral Schade and the Commanding Officer of Manley acted on

their own authority and immediately informed their superiors

of their actions, as specified in Navy Regulations.[49] The

episode also illustrates why Navy commanders have such broad

authority. U.S. officials in Washington were incapable of

staying abreast of a rapidly changing political situation.

Knowing only that rebel leaders had broadcast a warning

---

[48]Ibid; "Four to be Hanged in Zanzibar," Mombasa Times
(Mombasa, Kenya), January 15, 1964, p. 1.

[49]Manley received commendations from the Chairman of
the Joint Chiefs of Staff, Chief of Naval Operations,
Commander in Chief U.S. Forces Europe, Commander in Chief
U.S. Naval Forces Europe, and Commander Middle East Force
for "outstanding performance, vigilance, and prompt and
correct reactions to unusual conditions." See Chief of
Naval Operations message, CNO 081628Z FEB 64, February 8,
1964 (Unclassified, copy provided to author by Captain
Murray); Commander Middle East Force message, COMIDEASTFOR
170205Z FEB 64, February 17, 1964 (Unclassified, copy
provided to author by Captain Murray); Ship's History, USS
Manley (DD 940), 1964 (Ships History Branch, Naval
Historical Center, Washington, DC). Manley's Executive
Officer was personally commended for his role in the
evacuation. See James R. Ruchti, First Secretary, U.S.
Embassy, Nairobi, Kenya, letter to Commanding Officer USS
Manley (DD 940), March 17, 1964 (provided to author by
Captain Murray); "Three Manley Crewmen Commended for
Heroism," Charleston Evening Post (Charleston, SC), December
10, 1964, p. 10A.

against interference, but not knowing that Manley's Executive Officer and the American Charge d'Affairs were in contact with the rebels and had gained their permission for an evacuation, Washington prematurely cancelled the mission. That the situation was indeed serious enough to warrant immediate evacuation is shown by what happened to the American Charge d'Affairs four days later: Picard was arrested at gun point and expelled from the island due to U.S. refusal to recognize the new government.[50] The situation could have been much more serious with sixty-one potential hostages on the island, as the United States would learn later in Iran.

Allison's original organizational process model would seriously misconstrue both of these incidents. The model does not account for the substantial authority delegated to Navy commanders, including authority to disregard orders when warranted by unanticipated extraordinary circumstances and lack of immediate communication with higher authority. In both the 1953 Sea of Japan and 1964 Zanzibar cases, Allison's organizational process model would view the commanders as mindlessly carrying out pre-established organizational routines routines regardless of the desires of higher authorities--missing the crucial points that in the 1953 case there was ambiguity as to how close the search

[50] "Zanzibar Regime Seizes U.S. Consul at Gunpoint," New York Times, January 17, 1964, p. 1; Davids, p. 279.

and rescue force could approach the Soviet Union, and that

in the 1964 case authorities in Washington lacked sufficient

information to effectively control the operation. Admiral

Schindler allowed his superiors to exercise control by

negation when he sent the message stating his intention, and

in the absence of a timely response acted with prudence and

kept his ships clear of Soviet territorial waters. Rear

Admiral Schade disregarded orders issued by the Secretary of

Defense via the military chain of command in order to

respond to the evacuation request, and Manley carried out

the evacuation with caution to avoid incidents in a volatile

situation. Neither organizational routines nor evasion of

civilian control was a factor in these two cases.[51]

---

[51]The behavior of Navy and Marine Corps commanders
during the 1958 landings in Lebanon was much different from
that of Navy commanders in the Zanzibar episode. In the
Lebanon case, Navy and Marine officers twice refused
requests from the American ambassador to modify their plans
for the landing, which could be cited as an example of the
organization process model at work. The first request was
that the Marines not be landed over the beach to seize the
airport, that they be kept aboard ship and brought into the
harbor. This request was made after the first battalion of
Marines was ashore and deployed, which would have reqired
lengthy backloading of the men and their equipment.
Additionally, President Eisenhower had already announced the
landing to the world. This request simply came too late to
be executed effectively. The second request, made about
thirty minutes after the first, was for a company of Marines
to be sent from the airport to the presidential palace to
guard against a possible coup. Note that in making this
request Lebanese officials and the U.S. ambassdor completely
reversed their position from the earlier request. This
request was refused because the Marine commander thought
that the small force would have been in an exposed and
vulnerable position, and cut off from the main force. He
was also concerned that he would not have sufficient troops

Summary

In summary, the military chain of command, from the President down to the lowest levels, is founded on the principle of delegating control while retaining command. Tension between delegation and control is always present in the military chain of command. Pressures toward centralized control are driven by the complexity of modern warfare, fear of nuclear war, and efforts to exploit the force multiplier effect. Pressures toward decentralized control are driven by severe constraints on the ability of top-level authorities to effectively control tactical operations, and by the advantages gained by granting the on-scene commander flexibility to exercise initiative. Improvements in communications and information processing systems are unlikely to reduce the need for delegation of control.

The methods of exercising control cover a "tightness of control" spectrum ranging from very tight to very loose control. Toward the tight end of the spectrum are positive

---

to secure the airport and landing area until the second battalion had deen landed. The Marine commander's caution appears to have been warranted given that when the first column of Marines attempted to leave the airport the next day, they met Lebanese tanks and artillery massed on the road into Beirut, almost causing an armed clash. Thus, although even the Marine commander ashore felt his orders were excessively rigid, valid political and military considerations were at least as important as organizational routines in determining the decsions that were made. See Robert, McClintock, "The American Landing in Lebanon," U.S. Naval Institute _Proceedings_ 88 (October 1962): 65-79; Hadd, pp. 81-89.

direct control, and direct control by negation. Toward the loose end of the spectrum are monitored delegated control and autonomous delegated control. The United States armed forces rely on a flexible combination of direct and delegated control. Certain of the methods of control can be used in conjunction, and forces can be rapidly shifted from one method to another as the situation warrants. A distinction between initiatory actions and contingent responses is used to delegate authority to take certain actions, while withholding authority to take others. Navy command and control doctrine is unique in the scope of the authority granted to commanding officers, which even allows them to act contrary to orders when the situation warrants.

These are the basic concepts on which military command and control procedures are based. Emphasis is on delegation of discretionary decisionmaking authority in conjunction with appropriate methods of control. In addition to direct control via communications links, commanders can exercise indirect control of subordinates even after having delegated them substantial autonomy. The mechanisms of indirect control are the subject of the next section.

Mechanisms of Indirect Control

When a military commander delegates control of operational forces, he does not relinquish all control of those forces to his subordinate. In most cases, he retains

a certain amount of direct control, which can vary widely in tightness. Additionally, the commander has at his disposal various mechanisms of indirect control. Mechanisms of indirect control are orders, instructions, or detailed guidance issued to a commander prior to the start of a mission in order to ensure that the operational decisions he makes support the objectives and intentions of his superiors. Such instructions can range from being very detailed and specific to very general in nature. As the method of control being used moves across the "tightness of control" spectrum from tight to loose--that is, as the subordinate is granted increasing freedom from direct control--the importance of the mechanisms of indirect control increases. When a subordinate is operating under autonomous delegated control, with no direct communications links at all, the mechanisms of indirect control are the only means of control available.

There are five principle mechanisms of delegated control: the alert system, standing orders, mission orders, contingency plans, and rules of engagement. The first four will be discussed in this section. Rules of engagement will be discussed separately in the next section. The focus of attention in the following discussion will be on how the mechanisms are supposed to function in principle, rather than on hw they actually function in practice, which will be addressed in the case studies in Chapters VII and VIII.

The Alert System

The U.S. alert system is based on five levels of Defense Readiness Condition (DEFCON), ranging from normal peacetime readiness (DEFCON 5 and 4) to wartime readiness (DEFCON 1). The DEFCON system defines the overall framework for controlling the readiness of U.S. forces, providing a uniform system for all operational commands. Within this framework, following guidance from the Joint Chiefs of Staff (JCS), individual commands formulate alert procedures and readiness postures applicable to their forces. The system is highly flexible, allowing different major commands to be placed at different DEFCON levels as the world situation warrants. It is not unusual for U.S. forces in different parts of the world to be at different DEFCON levels.[52]

Much of the detailed guidance for operational forces is not part of the alert system per se, it is included in standing orders and contingency plans activated as higher levels of DEFCON are declared. Thus, the primary impact of the alert system is that it activates a wide range of operational guidance contained in previously prepared standing orders and contingency plans.

---

[52]Joseph J. Kruzel, "Military Alerts and Diplomatic Signals," in Ellen P. Stern, ed., The Limits of Military Intervention (Beverly Hills: Sage, 1977), pp. 83-99; Bruce G. Blair, "Alerting in Crisis and Conventional War," in Ashton B. Carter, John D. Steinbruner, and Charles A. Zraket, eds., Managing Nuclear Operations (Washington, DC: Brookings Institution, 1987), pp. 75-120; Sagan, pp. 100-102.

The alert system serves as a mechanism of indirect control by ensuring a uniform response to an order to increase (or decrease) readiness. A commander need not issue detailed orders to every subordinate command specifying the measures he wants them to take. The measures required for each level of readiness are spelled out in their standing orders, all the commander needs to do is state the level he desires and then monitor the messages that come in reporting attainment of the specified readiness condition. This frees the commander from having to directly manage details of implementing the alert, leaving him free to concentrate on assessing the situation and planning subsequent military operations.

An important feature of the alert system is that certain military commanders are delegated authority to increase the readiness of their forces independent of the DEFCON set by the JCS. They must maintain the minimum readiness level set by JCS, but can place their forces at a higher condition of readiness if warranted by the particular threat facing their commands. The Commanders in Chief of the unified and certain of the specified commands--such as the Atlantic Command, Pacific Command, U.S. European Command, and Strategic Air Command--have authority to increase the DEFCON of their forces independently of the worldwide DEFCON. They are required to immediately report such an action to the JCS and, time permitting, would normally confer with JCS before

changing the DEFCON level of their forces. The Commanders in Chief can also select from among various readiness postures--tailored for different types of threats--within a given DEFCON level.[53] Lower level commanders (who do not have authority to order changes in DEFCON) can also increase the readiness of their forces independent of the worldwide or theater DEFCON level. For example, the commanding officer of a Navy ship can place his crew at Condition I ("general quarters," when the crew is at battle stations) on his own authority without regard to DEFCON. A Navy battle group or fleet commander can place his entire force in an increased readiness status on his own authority. Similar procedures exist throughout the armed forces.

Standing Orders

Standing orders are detailed guidance on operational procedures prepared on a routine basis during peacetime. Although they are revised periodically, the intent is that they provide stable guidance, thereby minimizing uncertainty over operational procedures and facilitating the exercise of delegated control. Standing orders fall into four general categories: doctrinal publications, operations orders, operations plans, and long-range schedules.

---

[53]JCS SM-833-59, August 25, 1959; JCS 1968/84, Record Group 218, JCS Records, National Archives, Washington, DC. Also see Blair, "Alerting in Crisis and Conventional War," pp. 114-117; Sagan, pp. 134-135.

Doctrinal publications define strategic principles, standard tactics, and routine operational procedures. U.S. Army Field Manuals (FMs), U.S. Navy Naval Warfare Publications (NWPs), and NATO Allied Tactical Publications (ATPs) are examples of doctrinal publications.[54] Doctrine serves as a mechanism of indirect control by allowing a commander to issue a very brief order directing that a particular tactic or procedure be executed without having to specify all the details of the actions every unit is to take. Doctrine can be taught and rehearsed in peacetime, thus reducing confusion arising from unfamiliar procedures. Certain procedures require amplifying information for coordination purposes, such as direction of movement and timing of actions, but the prior formulation of a doctrine for carrying out the action still facilitates controlling it. Doctrine also reduces communications up the chain of command by reducing the amount of details that must be included in situation reports in order to explain the progress of an operation.

A common misconception about doctrine is that it specifies only a single tactic to be used in each situation. This misconception leads to the view that a military

---

[54]For a description of U.S. Navy doctrinal publications, see Commander George Galdorisi, "The Quiet Revolution," U.S. Naval Institute Proceedings 112 (April 1986): 42-43; "Surface Tactical Doctrine," Surface Warfare 10 (September/October 1985): 15-17.

organization's repertoire is limited to a small range of standard operating procedures, and consequently to overemphasis of the organizational process model for analysis of military organizations. In fact, doctrinal publications typically define a range of options for any given situation, and provide criteria for selecting among them. The complexity of doctrine and the range of options it encompasses has increased in direct proportion with the complexity of warfare. The greater the variety of sensors, platforms, and weapons at a commander's disposal, the greater are his options in any given situation.

The easy part of tactical training is teaching standard tactical procedures, the difficult part is teaching tactical decisionmakers how to select the most appropriate option, or combination of options, for a particular situation. Most doctrinal publications explicitly recognize that not every tactical situation can be anticipated, and that the standard procedures they contain should be used as building blocks for constructing an appropriate plan of action.[55] Command and control would not be as complex as it is if doctrine were in fact as simple as it is often incorrectly portrayed.

[55]Lieutenant Christopher H. Johnson, "Tactics," U.S. Naval Institute Proceedings 104 (October 1978): 37-43; Lieutenant Commander Miles A. Libbey, III, "Time Out for Tactics," U.S. Naval Institute Proceedings 104 (January 1979): 52-57; Captain Robert C. Powers, "The Return of Tactical Thought," U.S. Naval Institute Proceedings 109 (June 1983): 21-27.

Standing operations orders (OPORDs) are issued by operational commands to provide guidance for the routine peacetime operations conducted by their forces. The term "operations order" is a Navy term, but the other services have equivalent documents. A standing operations order typically defines command relationships, communications channels, logistics procedures, and other such routine information. It serves as a mechanism of indirect control by relieving a commander of having to repeatedly issue the same orders to cover repetitive routine situations. An important function of standing operations orders is to define the scope of decisionmaking authority delegated to subordinate commanders on a routine basis. Standing operations orders are a good example of hierarchical controls or rule-governed delegation of decisionmaking.

An operations plan (OPLAN) is a comprehensive set of plans for the conduct of a specific operation. Standing operational plans are those that have been activated for execution of particular peacetime operations. They serve a function similar to that of standing operations orders, but are limited in scope to a single operation. The single operation could well consist of a prolonged series of repetitive missions, such as reconnaissance flights. An operations plan includes the objective of the operation, the forces assigned to it, command relationships, communications channels, doctrinal guidance, intelligence procedures,

logistics, and a multitude of other information as appropriate. An operations plan serves as a mechanism of indirect control by relieving a commander of having to exercise direct control over the routine aspects of an operation. An important function of standing operations plans is to define the scope of decisionmaking authority delegated to subordinate commanders on a routine basis. Standing operations orders are another good example of hierarchical controls or rule-governed delegation of decisionmaking.

Long-range schedules are prepared for such recurring activities as surveillance missions, major maintenance periods for ships and aircraft, ship deployments, and exercises. Long-range schedules are often "nested," with shorter-term schedules filling in the details of longer-range schedules. Long-range schedules serve as a mechanism of indirect control by allowing a commander to approve a large number of recurring routine operations at one time, leaving the details of planning and executing each one to subordinates.

Long-range schedules are an important means of indirect presidential control of the U.S. armed forces. Long-range schedules function on the principle of control by negation: Once a schedule is approved, the operations listed in it are authorized for execution unless specifically cancelled. Schedules for certain operations that are especially

sensitive or visible, such as reconnaissance missions and major exercises, are reviewed by the Secretary of Defense and National Security Advisor after the long range schedule is approved (in some cases prior to specific operations). They review the schedules to ensure that the operations are still appropriate and will not interfere with U.S. foreign policy. If questions arise concerning the political implications of a particular operation, it may be cancelled or referred to the President for a final decision. Most long-range schedules for routine evolutions are not subjected to this close scrutiny after they are approved. Long-range schedules are subject to frequent changes due to the fickle nature of world politics, but they nonetheless serve important planning and control functions.

Mission Orders

Mission orders include letters of intent (LOIs), operations plans or operations orders issued for a specific short-term operation, and various other types of orders used to initiate routine and non-routine operations. Mission orders are important when the monitored delegated control method is used, and crucial when the autonomous delegated control method is used. They are less important when one of the direct control methods is used. Mission orders can range from being very detailed and specific to being very brief and general. At a minimum, a mission order includes

the objective of the operation, the forces assigned to it, the identity of the commander, and the time frame for the operation. But a mission order can include the same amount of detail as (and in fact be) an operations plan.

Mission orders serve as a mechanism of indirect control by relieving a commander of having to exercise direct control over the details of an operation's execution. An important function of mission orders is to define the scope of decisionmaking authority delegated to subordinate commanders. A mission order can specify which decisions must be referred to higher authority and which decisions the subordinate commander is authorized make himself. Thus, mission orders provide a means of allowing a commander to focus his attention on the most important decisions that come up in the course of an operation without being distracted by routine matters. This in turn helps to prevent communications channels and decisionmakers from becoming overloaded.

Contingency Plans

Contingency plans are those operations plans (OPLANs) prepared in advance for execution in the circumstances specified in the plans.[56] Contingency plans are commonly

---

[56]Note that operations plans can serve as contingency plans, mission orders, and standing orders. An individual OPLAN can transition from being a contingency plan (prior to execution), to being a mission order (upon execution), to being a standing order (for a long-term operation not requiring direct control).

prepared for crisis and peacetime emergency scenarios, various limited war scenarios, and general war scenarios (the last two types are often collectively referred to as "war plans"). Contingency plans serve as a mechanism of indirect control by allowing a commander to rapidly issue a single order to execute an operation that he and his staff have had time to prepare in detail ahead of time. Contingency plans are distributed in advance, eliminating the burden of having to issue a large volume of orders when a decision in made to carry out the operation. The only direct orders that are needed are last-minute revisions to the contingency plan and the mission order directing that it be executed as modified. Once a contingency plan is executed, it serves the same functions described above for operations plans and operations orders.

The single most important U.S. contingency plan is the Single Integrated Operational Plan (SIOP) for the employment of U.S. strategic nuclear forces.[57] Contingency plans are often designed to be mutually supportive with other

[57]See Henry S. Rowen, "Formulating Strategic Doctrine," in Commission on the Organization of the Government for the Conduct of Foreign Policy, Report on the Organization of the Government for the Conduct of Foreign Policy, Volume 4, Appendix K, Part III (Washington, DC: U.S. Government Printing Office, June 1975), pp. 219-34; Desmond Ball, "Targeting for Strategic Deterrence," Adelphi Papers No. 185 (London: Interrnational Institute for Strategic Studies, Summer 1983), pp. 8-25; Scott D. Sagan, "SIOP-62: The Nuclear War Plan Briefing to President Kennedy," International Security 12 (Summer 1987): 22-51.

contingency plans, as well as capable of being executed individually. For example, the overall U.S. strategic objective of attempting to prevent a general war with the Soviet Union from escalating to a strategic nuclear exchange requires that the war plans of the unified commands be capable of execution independently as well as in conjunction with the SIOP. An important feature of the U.S. alert system is that increasing the readiness condition activates preparations to execute designated contingency plans applicable to the conflict at hand. This type of preparation is closely linked to the alert system because the effectiveness of an increase in DEFCON is much greater when the forces have a specific mission that they are preparing to execute than when they simply increase their readiness without being assigned a specific mission.

There are two types of contingency planning: routine and crisis. Routine contingency planning takes place on a continuing basis in peacetime to prepare plans for the most likely and most dangerous situations that can be anticipated. The Joint Chiefs of Staff and the commanders in chief of the unified commands have primary responsibility for deciding what contingency plans are written. The National Security Council has only a very small role in routine contingency planning. There is limited liaison between the State Department and the Joint Chiefs in certain types of contingency planning, such as for evacuation of

embassy personnel and their families in crises.[58] The

difficulties in routine contingency planning are

anticipating the possible crises for which planning should

be done, defining specific scenarios for the use of force in

each situation, predicting the forces that will be available

for executing various options, and assessing the political

feasibility of different options.

Crisis contingency planning consists of revising

existing plans or formulating new plans for a range of

military operations after a crisis has arisen. The National

Security Council plays a significant role in determining

which contingency plans are updated or prepared during

crisis contingency planning. Normally, the Joint Staff and

the unified command responsible for the area in which the

crisis is located begin crisis contingency planning as soon

as indications are received that a crisis is about to

break.[59] Although many of the difficulties in routine

---

[58]William A. Cockell, Jr., Director, Defense Policy
Directorate, National Security Council, interview by author,
February 11, 1988; Ronald St. Martin, Director, Office of
Program Integration, Office of the Secretary of Defense, and
formerly Director, Crisis Management Center, National
Security Council, interview by author, February 8, 1988;
Colonel Charles J. Bauer, "Military Crisis Management at the
National Level," Military Review 55 (August 1975): 3-15;
"Understanding Military Contingency Planning," Military
Review 61 (July 1981): 33-43; Admiral Robert L. Dennison,
"Reminiscences of Admiral Robert L. Dennison, U.S. Navy
(Retired)," Oral History Program, U.S. Naval Institute,
August 1975, pp. 296-297; Gallery, pp. 27-28, 85-86.

[59]Cockell interview; St. Martin interview; Hayward, p.
261.

contingency planning are eliminated because planning is being done for a specific crisis, new problems arise due time pressures on the planning process and the ambiguity and confusion that always surround a crisis.

The Joint Chiefs of Staff are often described as incapable or unwilling to provide a broad range of military options for dealing with a crisis, but this does not appear to be accurate. When asked about this, the consensus among present and former National Security Council staff members, Department of Defense officials, and Joint Staff officers is that the JCS will provide a range of options when directed to do so. They always have (and press for) a preferred option, which typically entails decisive use of superior force so as to reduce the risk of defeat and to deter escalation. It is thus not unusual for the option preferred by the JCS to be viewed as politically infeasible by civilian authorities. The JCS also tend to resist civilian involvement in the details of operational planning once the objectives and basic parameters of an operation have been defined.[60]

Another aspect of military contingency planning for crises is that the Joint Chiefs and the Commanders in Chief

---

[60]Richard K. Betts, Soldiers, Statesmen, and Cold War Crises (Cambridge, MA: Harvard University Press, 1977), pp.160-161; Marshall Brement, "Civilian-Military Relations in the Context of National Security Policymaking," Naval War College Review 41 (Winter 1988):28-29; Cockell interview; St. Martin interview.

of the unified commands can execute a wide range of
preparatory actions as soon as initial indications are
received that a crisis is breaking. These include
activating special crisis management staffs, informing
subordinate commanders of the situation, assembling
personnel with specialized training and experience in the
area, canceling or modifying routine operations that might
exacerbate the crisis or delay a military response,
increasing surveillance in the area, clearing communications
channels and setting up special channels, increasing the
readiness of units iden-tified in contingency plans, and
even deploying certain units that might be needed on short
notice. Many such actions can be executed without prior
approval of the President, though he is normally informed of
them via the Secretary of Defense or the National Security
Advisor as soon as possible after they are taken.[61]

---

[61]Hayward, p. 261; briefings for author at Navy Command
Center, National Military Command Center, and National
Security Council, February 1988. Admiral Arleigh Burke,
Chief of Naval Operations 1955-1961, was particularly adept
at quietly alerting Navy units and readying them for
potential crises. It was due to such actions by the CNO
that the Navy was able to conduct the 1958 landing in
Lebanon on short notice. See Admiral Arleigh A. Burke, "The
Lebanon Crisis," in Arnold A. Shapack, ed., Proceedings of
the Naval History Symposium (Annapolis, MD: U.S. Naval
Academy, April 27-28, 1973), pp. 72-75. On military crisis
management procedures, see Head, Short and McFarlane, pp. 64-
67; Joint Chiefs of Staff, "Crisis Staffing Procedures of
the Joint Chiefs of Staff," SM-481-83 (1983); Office of the
Chief of Naval Operations, "Navy Crisis Management
Organization," OPNAVINST 1601.7G, January 8, 1988. Also see
Blair, "Alerting in Crisis and Conventional War," pp. 114-
117; Sagan, pp. 134-135.

Preparation of contingency plans for retaliatory attacks is a routine part of military planning for peacetime and crisis operations in which hostilities could occur. As a general rule, authority to order retaliatory attacks is not delegated to military commanders, it is closely held by the President. Contingency planning for such attacks makes an important contribution to the effectiveness of peacetime and crisis military operations by providing the President with military options for a prompt response to provocations. In contingency planning for retaliation, as in all other types of routine contingency planning, preparation of a plan is not an indication of intent to execute the plan.

Opinions vary widely as to the value of routine contingency planning for crisis management. One study found that from 1946 to 1975 (41 cases) no appropriate contingency plans were available in 58.5% of the cases, and that the available contingency plans were inadequate in another 24.4% of the cases.[62] Although these would appear to be grim statistics, the study did not address the important value that contingency plans can have even when inappropriate or requiring modification (this will be discussed below). Contingency planning problems are caused by the difficulty of predicting where crises will arise, the specific

---

[60]Leo Hazelwood, John J. Hayes, and James R. Brownell, Jr., "Planning for Problems of Crisis Management," _International Studies Quarterly_ 21 (March 1977): 93.

political-military circumstances that will surround a crisis, how the President will react to a crisis, and the options the President will view as politically feasible. These factors impose inherent constraints on the ability of routine contingency planning to provide plans that are ideally suited for management of particular crises.

Although the plans that are produced by routine contingency planning may not be appropriate or adequate for the specific crises that arise, the planning process itself can make an important contribution to crisis management. Rear Admiral J.C. Wylie, former Deputy Commander in Chief U.S. Naval Forces Europe and an experienced Navy planner, has explained the value of contingency planning: "Contingency planning rarely fits unexpected situations. Reliance must be placed on uniformed officers. The most important benefit of contingency planning is that it trains planners, which is important."[63] Admiarl Horacio Rivero, Jr., former Commander in Chief Soutiiern Europe and Vice Chief of Naval Operations, has offered similar views: "You will learn, if you didn't know it before, that available plans are never carried out as written. They have to be modified to fit the particular circumstances at the time, and you have to do a considerable amount of improvising outside the plans. However, the previously prepared plan is

---

[63]Rear Admiral J.C. Wylie, letter to author, April 13, 1988.

essential to serve as a basis for modifications and improvements, and to make sure that you haven't forgotten anything under the pressure of time."[64] The United States conducted considerable contingency planning for military action against Cuba in the year prior to the 1962 Cuban Missile Crisis. Although none of the plans were executed as written, Admiral Afred G. Ward, commander of the Quarantine Force during the crisis, argues that the planning process contributed to U.S. readiness when the Soviet missiles were discovered: "This planning stood us in good stead at the time of the Cuban confrontation later, in which President Kennedy decided to take firm action to stop this movement of equipment, of goods and supplies, into Cuba. We were as ready as any nation has ever been to win a military victory in the period of October 1962."[65]

The comments offered by Admirals Wylie, Rivero and Ward reveal three reasons for the value of contingency planning. First, contingency planning educates the staff officers that prepare the plans and the commanders that review them on the characteristics of the area in which the operations will be conducted. Second, the operational and logistical problems

---

[64]Admiarl Horacio Rivero, Jr., letter to author, March 10, 1988.

[65]Admiral Afred G. Ward, "Reminiscences of Admiral Afred G. Ward, U.S. Navy (Retired)," (Annapolis, MD: U.S. Naval Institute, Oral History Program), p. 171. See Chapter VII for a detailed description of U.S. contingency planning prior to the Cuban Missile Crisis.

likely to be encountered are identified in the planning

process, even though the proposed solutions may not be

deemed appropriate for the specific crisis that arises.

Third, a contingency plan provides a baseline or starting

point for further planning after a crisis arises. It is

almost always easier to modify an existing plan than to

create a new plan from scratch, and even when a plan

requires extensive modification, it usually contains much

valuable information.[66] Thus, the contingency planning

process itself can contribute to crisis management.

The disadvantage of contingency planning is that the

existence of a plan can preclude other options that might be

superior for dealing with the crisis. A perception can

arise that there is not sufficient time to prepare

alternative plans. Support for an existing plan can

restrict the search for alternatives, or can create a bias

against alternative plans that have not been staffed as well

due to time constraints. The officers who prepared the

original plans sometimes resist modifications to them out of

pride in authorship.[67] However, officials who have had

---

[66]Also see Philip A. Odeen, "Organizing for National Security," International Security 5 (Summer 1980): 118; John M. Collins, U.S. Defense Planning: A Critique (Boulder, CO: Westview Press, 1982), pp. 11, 158-159.

[67]Ole R. Holsti, Crisis Escalation War (Montreal: McGill-Queens University Press, 1972), pp. 215, 235-237; Betts, pp. 154-156; Lebow, pp. 232-237; Cockell interview; St. Martin interview.

first-hand experience with crisis contingency planning
generally agree that momentum for an existing plan and pride
of authorship tend not to be serious problems if the
civilian leadership insists on a range of options and on
tailoring a plan that suits its objectives.[68] On balance,
then, routine contingency planning is far superior to doing
no planning at all, but usually cannot eliminate the need
for further planning after after a crisis erupts and can
create pressures that hamper the tailoring of military
options to meet crisis management objectives.

Summary

In summary, this section has reviewed four of the five
primary mechanisms of indirect control: the alert system,
standing orders, mission orders, and contingency plans. The
mechanisms of indirect control relieve higher authorities of
the burden of having to closely monitor the details of
military operations--a burden that can quickly exceed their
information processing and decisionmaking capabilities when
large-s ale operations are being conducted in a fast-paced
political-military environment. Relieved of this burden,
top-level authorities are better able to concentrate on
monitoring the overall political-strategic situation, formu-
lating and revising their strategy for dealing with the

[68]Betts, pp. 160-161; Odeen, p. 118; Cockell interview;
St. Martin interview.

confrontation, and coordinating the overall execution of military operations so that they support that strategy. The mechanisms of indirect control thus aid in striking an appropriate balance between autonomy and control in the execution of military operations, and, when used properly by national leaders, can contribute to crisis managment.

Thus far this discussion of the mechanisms of indirect control has focused on how they are supposed to work in principle. Neither the mechanisms themselves nor the decisionmakers that use them are are perfect. Many things can go wrong in the stress and confusion of crisis military operations. More importantly, there are inherent limits on the ability of these, or any, mechanisms to ensure that decisions made at one level are those that are most appropriate for the situation at another level. For example, national leaders could give tactical orders that are disastrous for the on-scene forces, or tactical commanders could take an action that seriously disrupts the crisis management efforts of national leaders. This problem is inherent because decisionmakers at the different levels are operating in different environments. They can develop much different threat perceptions, priorities of objectives, and expectations as to the future course of the crisis. How the mechanisms of indirect control perform in practice will be examined in Chapters VII and VIII, which present the case studies.

In addition to the four mechanisms of indirect control already discussed, there is a fifth mechanism: rules of engagement. Because rules of engagement are particularly important in crisis management, they will be discussed in greater detail in the next section.

Rules of Engagement

Rules of engagement are orders issued to define the circumstances in which the U.S. armed forces are authorized to use their weapons for defense against hostile forces in peacetime, and to specify the scope and level of violence of combat operations in wartime.[69] Rules of engagement serve as a mechanism of indirect control by allowing top-level authorities to specify policies on the use of force prior to situations in which direct control of the decision to use force is not possible. As Captain J. Ashley Roach has pointed out, rules of engagement are a tool for implementing top-level decisions on the use of force at the operational level, providing a means of ensuring that "national policy will be followed in wartime or sudden emergencies which do not allow time for communications between Washington and the

---

[69]J. Ashley Roach, "Rules of Engagement," Naval War College Review 36 (January-February 1983): 46-48. The official JCS definition is that they are "Directives issued by competent authority which delineate the circumstances and limitations under which United States forces will initiate and/or continue combat engagement with other forces encountered." See JCS Publication No. 1.

field."[70] In short, the purpose of rules of engagement is to provide guidance to operating forces from National Command Authorities, via the Joint Chiefs of Staff and the operational chain of command, on how to respond to threat of attack in peacetime, and on limitations on fighting in wartime.

Wartime rules of engagement place limits on military action when U.S. forces are engaged in an armed conflict.[71] Certain military options may be deemed undesirable in wartime due to escalation control, diplomatic, and humanitarian considerations. For example, an important

---

[70]Ibid, p. 47. Also see Rear Admiral J.R. Hill, Maritime Strategy for Medium Powers (Annapolis, MD: Naval Institute Press, 1986), p. 127.

[71]Roach, p. 49; U.S. Department of the Navy, The Commander's Guide to the Law of Naval Operations, Naval Warfare Publication No. 9 (Washington, DC: Office of the Chief of Naval Operations, July 1987), pp. 5-3, 5-4 (Cited hereafter as NWP 9). This publication was formerly titled Law of Naval Warfare, (Naval Warfare Information Publication 10-2). On the wartime rules of engagement used by U.S. forces during the Korean War, see Field, pp. 395-6; Malcolm W. Cagle and Frank A. Manson, The Sea War in Korea (Annapolis, MD: Naval Institute Press, 1957), pp. 224-5, 243-7, 445, 458; Robert F. Futrell, The United States Air Force in Korea, 1950-1963 (New York: Duell, Sloan and Pierce, 1961): 142, 208-11, 453. On the wartime rules of engagement used by U.S. forces during the Vietnam War, see U.S. Congress, Congressional Record, Vol. 121, Part 14 (Washington, DC: U.S. Government Printing Office, 1975), pp. 17551-17558. On the role of rules of engagement in the 1982 war between Britain and Argentina over the Falkland Islands, see Christopher Craig, "Fighting by the Rules," Naval War College Review 37 (May-June 1984): 23-27; Max Hastings and Simon Jenkins, The Battle for the Falklands (New York: W.W. Norton, 1983), pp. 82, 124, 137, 147-8. For a discussion of how rules of engagement change from peacetime to wartime, see Hill, pp. 127-9, 133-4, 142.

escalation control function of wartime rules of engagement is to prevent incidents with the military forces of non-belligerents. Wartime rules of engagement can also be used to prevent geographic expansion of a conflict when it is politically and diplomatically desirable to confine the fighting to a limited area (i.e., prohibitions against attacking the homeland when fighting at sea). Wartime rules of engagement allow military action under such circumstances only for self-defense--the adversary is forced to make the decision to escalate or expand the conflict.

Peacetime rules of engagement are founded on the right of self-defense as defined under international law and in U.S. Department of Defense directives.[72] Simply put, peacetime rules of engagement prohibit U.S. military commanders from shooting first in peacetime unless

---

[72]On United States policy, see NWP 9, pp. 4-1 to 4-5; U.S. Department of Defense, "DoD Law of War Program," Department of Defense Directive 5100.77, July 10, 1979; U.S. Department of the Army, Law of Land Warfare, Army Field Manual FM 27-10 (Washington, DC: U.S. Government Printing Office, 1956); U.S. Department of the Air Force, International Law--The Conduct of Armed Conflict and Air Operations, Air Force Pamphlet 110-31 (Washington, DC: U.S. Government Printing Office, 1977). Also see Lieutenant Commander Bruce Harlow, "The Legal Use of Force...Short of War," U.S. Naval Institute Proceedings 92 (November 1966): 88-98. On international legal principles, see Derek Bowett, "Reprisals Involving Recourse to Armed Force," American Journal of International Law 66 (January 1972): 1-36; Ian Brownlie, "The Use of Force in Self-Defense," British Year Book of International Law, 1961 (London: Oxford University Press, 1962), pp. 183-268; D.P. O'Connell, The Influence of Law on Sea Power (Manchester: Manchester University Press, 1975), pp. 70-84.

absolutely necessary for self-defense. Peacetime rules of engagement are intended to prevent unwanted military incidents and support crisis management. As George Bunn has pointed out, rules of engagement are intended "to restrain aggression, prevent the outbreak of hostilities, and to limit escalation if shooting starts."[73] Thus, peacetime rules of engagement are central to the problem of coordinating military policy with political and diplomatic objectives in a crisis.

There are two categories of peacetime rules of engagement: standing and special. Standing rules of engagement are written for routine peacetime operations. They are in effect at all times for the forces they cover. Special rules of engagement are issued to cover particularly sensitive situations, such as operations near a country openly hostile to the U.S. and operations during an international crisis. Special rules of engagement may replace or supplement standing rules of engagement, and may be be either more or less restrictive than standing rules of

---

[73]George Bunn, "International Law and the Use of Force in Peacetime: Do U.S. Ships Have to Take the First Hit?" Naval War College Review 39 (May-June 1986): 69. Also see Roach, pp.46-7; Norman Friedman, "The Rules of Engagement Issue," in E.F. Guertz, et al., NATO's Maritime Strategy: Issues and Developments (Washington, DC: Pergamon-Brassey's, 1987), pp. 23-4; Jacquelin K. Davis, et al., "NATO's Maritime Defenses," in Francis J. West, Jr., et al., Naval Forces and Western Security (Washington, DC: Pergamon-Brassey's, 1986), p. 46.

engagement, depending on the political-military circum-
stances. Examples of special rules of engagement include
the rules issued for the 1958 Marine landings in Lebanon,
the 1962 Cuban Missile Crisis, the August 1981 freedom of
navigation operations in the Gulf of Sidra, the 1983 Marine
peacekeeping force in Beirut, and the 1983 invasion of
Grenada.[74] A standardized format for ordering and modifying
rules of engagement is used throughout the U.S. armed forces
for ease and clarity when issuing special rules of
engagement and modifying standing rules of engagement.[75]
This system allows for ease of adapting rules of engagement
to changing political-military circumstances and to specific
U.S. foreign policy objectives in situations of increased
tensions.

The President, as commander in chief of the armed
forces, is the ι timate source of all rules of engagement.
At the top of the chain of command, overall guidance on

---

[74]See "Department of Defense Operations During the
Cuban Missile Crisis," Naval War College Review 32
(July/August 1979): 85; Admiral William H. Rowden, "Sixth
Fleet Operations: June 1981 to July 1983," in James G. March
and Roger Weissinger-Baylon, eds., Ambiguity and Command:
Organizational Perspectives on Military Decision Making
(Marchfield, MA: Pitman Publishing, 1986), pp. 271-5;
Metcalf, p. 281. Navy aircraft supporting the Marines
ashore in Lebanon from July to October 1958 were ordered
"not to return fire" when fired on by rebel forces. See USS
Essex (CVA 9), Ship's History 1958, Ships History Branch,
Naval Historical Center, Washington, DC.

[75]C.C. Pease, "Comment and Discussion," U.S. Naval
Institute Proceedings 108 (June 1982): 83; Roach, pp. 51.

rules of engagement is formulated by the JCS (with the
support of the Joint Staff and individual service chief
staffs). Civilian authorities are involved in the
formulation, review and approval of rules of engagement,
although in practice the level of their involvement varies
widely. Routine revisions to standing rules of engagement
receive little attention from civilian officials other than
the Secretary of Defense (and his aides that deal with such
matters). Civilian authorities become directly involved in
the formulation and review of special rules of engagement
and major revisions to standing rules of engagement. The
NSC interdepartmental group chaired by the Deputy National
Security Advisor and composed of top deputies from the
Department of State, Department of Defense, and JCS
(normally the Assistant to the Chairman)--currently known as
the Policy Review Group--reviews proposed rules of
engagement to ensure that they support overall presidential
policies. The role of this group is especially prominent in
crises. The National Security Advisor reviews important
revisions to rules of engagement, and submits revisions
involving relaxations of restrictions to the President for
approval.[76]

---

[76]NWP 9, p. 5-3; Roach, p. 51; Cockell interview; St.
Martin interview; briefings for author at Navy Command
Center, National Military Command Center, and National
Security Council.

Rules of engagement are promulgated via the operational chain of command, with increasing specificity at each successive level--reflecting the unique strategic and tactical circumstances of individual commands. These successive additions to the rules of engagement are not intended to modify the rules of engagement, their purpose is to tailor what is usually generalized guidance to specific circumstances. Commanders in chief of the unified commands and subordinate commanders under them must submit proposed revisions to their rules of engagement via the chain of command to the Joint Chiefs of Staff for approval. All rules of engagement promulgated by commanders with authority to promulgate their own rules are submitted to the JCS, and the Joint Staff maintains an up-to-date file of them. The JCS routinely sends copies of the rules of engagement to the NSC Staff, which keeps them on file for ready reference. The NSC has an individual on the Situation Support Staff charged with maintaining the NSC rules of engagement file and serving as the in-house expert on the rules currently in effect.[77] Thus, top-level military and civilian authorities directly involved in crisis management have ready access to all rules of engagement promulgated to U.S. forces.

---

[77]Ibid. For an excellent description of this process in action, see Department of Defense Commission on Beirut International Airport Terrorist Act of October 23, 1983, "Report of the DOD Commission on Beirut International Airport Terrorist Act, October 23, 1983," December 20, 1983, pp. 44-51. (Cited hereafter as Beirut Commission.)

Employment of rules of engagement as a method of indirect control entails a two-stage decision process. In the first stage, the Secretary of Defense, acting on behalf of the President, formulates rules of engagement that will support national political-diplomatic objectives and policies (the President normally reviews and makes the final decision on major or particularly sensitive revisions to the rules). In the second stage, the on-scene commander (the senior officer in command of the forces at the scene of a crisis) and the tactical decisionmakers under him (commanding officers of individual units and watch officers with authority to order the use of weapons) use the rules of engagement as guidance for making operational decisions on the use of force. Rules of engagement do not require that a commander attempt to consult with higher authority before taking action in self-defense. The rules exist specifically because commanders in the field or at sea may not have the means or sufficient time to contact higher authority. The fundamental objective of rules of engagement is for the operational decisions made by tactical commanders to support national objectives and policies as well as ensuring the defense of U.S. forces.

Rules of engagement are a central element in the flexible U.S. system of direct and delegated control. Rules of engagement are an important element in the guidance provided to the on-scene commander defining the scope of

tactical decisions he is authorized to make. Restrictive rules of engagement narrow the tactical decision-making authority of the on-scene commander, in effect imposing a greater degree of direct control on him. Permissive rules of engagement broaden his decision-making authority, in effect shifting him toward greater delegated control.

The authority to revise rules of engagement is itself an important issue in the balance between direct and delegated control. Reserving authority to revise rules of engagement to top-level authorities can have the same negative impact as attempting direct control of operations, while delegating authority to revise rules of engagement to lower levels can raise the same problems of coordinating national policies as delegated control. Commanders in the chain of command, including the on-scene commander, usually have limited authority to revise the rules of engagement, when such revisions do not result in a significant relaxation of the rules. Authority to make broader revisions to the rules of engagement, particularly to issue significantly more permissive rules, is reserved for top-level authorities. Rules of engagement are thus affected by the same tension between delegation and control that affects all other aspects of command and control.

Guidance from other sources of operational guidance is often incorrectly attributed to rules of engagement. For example, there is an important distinction between rules of

engagement, which govern how to handle to potentially
hostile forces, and measures to avoid mutual interference,
which govern how to avoid engagements with friendly forces
(these come under the category of standing orders as a
mechanism of indirect control). The two categories overlap
in that both address requirements for identifying
unidentified contacts. At times during the evolution of
rules of engagement, particularly during the Vietnam War
(when fighter pilots were required to visually identify air
targets in order to avoid firing on other U.S. planes),
measures to avoid mutual interference were included in the
rules, but this has subsequently been corrected. In some
warfare environments, such as in NATO, where friendly forces
from several countries may operate in the same battle area,
measures to avoid mutual interference are particularly
important, even overshadowing rules of engagement. However,
measures to avoid mutual interference do not have a
significant role in crisis management, so will not be
discussed further.

The remainder of this discussion will focus on
peacetime rules of engagement. Although wartime rules of
engagement are of great importance, particularly in the
study of limited war, they raise a different set of issues
than those of interest in crisis management. The topics
that will be examined are the reasons why rules of
engagement are needed, the history of rules of engagement,

the nature of peacetime naval rules of engagement, how the

decision to use force is made by an operational commander,

the impact of political-military context on decisions to use

force, the problem of rules of engagement being

misinterpreted, and the coordination of rules of engagement

among allies.[78]

The Need for Rules of Engagement

Rules of engagement are necessary for five reasons.

First, and foremost, commanders in the field or at sea may

not have the means or sufficient time to contact higher

authority. The speed of modern warfare causes the tactical

situation to change much faster than it can be explained to

higher authority, and the destructiveness of modern weapons

can make decisionmaking delays fatal. Rules of engagement

are a form of contingent response: action, in this case use

of force, can only be taken under specified conditions.

Initiatory actions, such as retaliatory attacks or

---

[78]This discussion of peacetime rules of engagement is
applicable to all of the U.S. armed forces. The basic
principles and concepts presented in the section on
peacetime naval rules of engagement are also applicable to
the other services, although the focus is on how they apply
to naval forces. Rules of engagement affect the day-to-day
operations of the Navy and Air Force much more than those of
the Army. This is because national borders tend to keep
ground forces separated, but U.S. ships on the high seas and
planes in international airspace are frequently in close
proximity to those of potential adversaries. Differences in
the three services' rules of engagement are primarily due to
differences in their command structures and warfare
environments, rather than differences in basic principles.

pre-emption in the absence of an imminent threat, are excluded. Thus, rules of engagement do not require that a commander attempt to consult with higher authority before taking action in self-defense.

The second reason why rules of engagement are necessary is that the lethality of modern weapons--particularly the anti-ship cruise missile--makes it exceedingly dangerous to take the first hit.[79] U.S. Navy ships and aircraft are generally authorized to take defensive action upon clear demonstration of hostile intent.[80] United States Navy Regulations, 1973 state that "The right of self-defense may arise in order to counter either the use of force or an immediate threat of the use of force."[81] In especially volatile situations, a clear demonstration of hostile intent may be limited to actual use of weapons by the adversary. This can lead to what Rear Admiral Hill calls the "concept of initial casualty": "This is to say that it may be necessary, in the opening rules of engagement, to accept the risk of a casualty before the relaxations necessary to allow prudent self-defense can be made."[82] Occasionally, the rules of engagement may prohibit use of force even when

---

[79]Friedman, pp. 32-3; O'Connell, pp. 81-2.

[80]Roach, pp. 49-50; Bunn, p. 69.

[81]United States Navy Regulations, 1973, p. 38. Also see NWP 9, p. 4-4.

[82]Hill, p. 128. Also see O'Connell, pp. 82-4.

fired upon, as was the case for Navy aircraft over Lebanon in 1958.[83] A primary function of rules of engagement is thus to define the actions and indicators that are to be used to determine that hostile intent is being demonstrated.

The third reason why rules of engagement are necessary is that not all nations that are potentially hostile to the U.S. present the same level of military threat to U.S. forces. U.S. forces legitimately need great leeway toward certain openly hostile and militarily unpredictable countries, particularly when they have attacked U.S. forces in the past. On the other hand, the U.S. has evolved fairly stable, tacit "rules of the game" in its military relationships with other countries, particularly the Soviet Union.[84] With such countries the threat is more predictable and greater care can be taken to avoid inadvertent incidents without unnecessarily risking U.S. forces. The interaction

---

[83]USS _Essex_ Ship's History.

[84]On the concept of "rules of the game" in Soviet-American relations, see James N. McConnell, "The 'Rules of the Game': A Theory on the Practice of Superpower Naval Diplomacy," in Bradford Dismukes and James McConnell, eds., _Soviet Naval Diplomacy_ (New York: Pergamon, 1979), pp. 240-280; Robert Legvold, "The Super-Rivals: Conflict in the Third World," _Foreign Affairs_ 57 (Spring 1979): 755-778; Raymond Cohen, _International Politics: The Rules of the Game_ (London: Longman, 1981); Joanne Gowa and Nils H. Wessel, _Ground Rules: Soviet and American Involvement in Regional Conflicts_ (Philadelphia: Foreign Policy research Institute, 1982); and Neil Matheson, The 'Rules of the Game' of Superpower Military Intervention in the Third World, 1975-1980 (Washington, DC: University Press of America, 1982), pp. 99-117.

of U.S. and Soviet naval forces is also regulated by the
international "rules of the road" governing the safe
navigation of ships at sea, supplemented by the Soviet-
American Agreement on the Prevention of Incidents On and
Over the High Seas signed in 1972. The rules of engagement
reflect this wide range in the stability and predictability
potential military threats, providing more permissive rules
when the danger is greater and more restrictive rules when
interactions are better regulated.

The fourth reason why rules of engagement are necessary
is to ensure that in responding to a hostile act or hostile
intent, U.S. forces adhere to the international legal
principles of proportional force and minimum force. The
principle of proportional force requires that the force used
in self-defense be proportional to the force used in the
hostile act or threatened when hostile intent was shown.
The principle of minimum force requires that the level of
force used in response to a hostile act or hostile intent be
limited to the minimum necessary to prevent the threat of
further attack.[85] Rules of engagement provide guidance on
the types of defensive actions that are authorized under
various circumstances.

The fifth reason why rules of engagement are necessary
is that U.S. forces can be tasked by the President to defend

---

[85]Bunn, pp. 73-74; O'Connell, p. 171; Roach, p. 50.

civilian U.S. vessels and U.S. citizens ashore overseas, the military forces and civilian vessels of allies and friendly nations, and the territory of allies or friendly nations. Rules of engagement are used to spell out when defensive action may be taken in such circumstances.[86]

History of Rules of Engagement

The United States Navy has had almost two centuries of experience with political limitations on the use of force due to its role as an implement of foreign policy in the nineteenth century. For example, when Commodore Matthew C. Perry was dispatched to negotiate a commerce treaty with Japan in 1853, the Secretary of State warned: "He will bear in mind that, as the President has no power to declare war, his mission is necessarily of a pacific character, and will not resort to force unless in self-defense in the protection of the vessels and crews under his command, or to resent an act of personal violence offered to himself or to one of his officers."[87] Provisions similar to this were not uncommon in the sailing orders given to American captains before they departed for distant stations.

---

[86]Bunn, p. 69.

[87]Quoted in Commander Dennis R. Neutze, "Bluejacket Diplomacy: A Juridical Examination of the Use of Naval Forces in Support of United States Foreign Policy," JAG Journal 32 (Summer 1982): 111.

Prior to World War II, there was little need for rules of engagement other than for Navy ships on diplomatic missions. The United States was far removed from potential enemies and its forces were rarely in contact with those of potential adversaries. All this changed with the advent of long-range aircraft and the growth of American global security commitments. As the Cold War with the Soviet Union intensified, confrontations between the superpowers became more frequent and dangerous. This was starkly apparent in the 1948 Berlin crisis, when Soviet closing of ground access to the city and harassment of supply flights threatened to cause armed clashes. In the late 1940s the Soviets began shooting down American aircraft patrolling the periphery of Soviet airspace, and there were air battles between American and Soviet planes during the Korean War.[88] There thus arose a need for guidance on the use of force in peacetime.

---

[88] On the 1948 Berlin Crisis, see Lucius D. Clay, Decision in Germany (Garden City, NY: Doubleday, 1950); W. Phillips Davison, The Berlin Blockade (Princeton: Princeton University Press, 1958); Frank Howley, Berlin Command (New York: Punam's, 1950). There were two serious incidents between American and Soviet aircraft during the Korean War, both involving Soviet planes threatening U.S. Navy ships, as well as several other lesser incidents. See Cagle and Manson, pp. 469-74; Field, pp. 167-9, 440-1; Futrell, pp. 142, 567. The first incident in which a Soviet fighter attacked an American plane off the Soviet coast occurred on October 15, 1945, less than two months after Japan surrendered. Between 1945 and 1950 there were at least nine instances of Soviet or Warsaw Pact fighters attacking American or British planes. See Office of the Chief of Naval Operations (OP-09B91R4), "Soviet Attacks on Western Planes," memorandum dated July 15, 1960 (Operational Archives, Naval Historical Center, Washington, DC).

In 1950 President Truman approved a comprehensive policy statement on interception of aircraft in United States airspace, the first such policy issued since the end of World War II. This initial guidance was later assessed to be overly restrictive, and was replaced by a revised interception policy in 1952. Although the term rules of engagement per se had not yet officially entered the military vocabulary, these two presidential directives constitute the origins of Unites States rules of engagement.

The U.S. Air Force was the first service to begin using the term rules of engagement to describe intercept and engagement policy, and in the fall of 1952 began using the term in the joint planning arena. In 1958 the Joint Chiefs of Staff officially adopted the term rules of engagement and defined them as "that body of authoritative law, instructions, policies, directives, measures, plans or decisions which authorize, restrict or describe the circumstances under which, and at times the means with which, U.S. forces will or may initially engage enemy forces and the extent to which the engagement will be carried."[89] The unified commands shifted over to the new term at about the same time. The Navy, which had long used the term "measures for self-preservation in peacetime" to describe the same idea, was slow to convert to the term rules of

---

[89] See JCS Publication No. 1.

engagement. In fact, the term was not widely used in the fleet until the Vietnam War, when rules of engagement became a major factor in shaping combat operations.

All of the major concepts upon which United States rules of engagement are based were adopted during the 1950s. The Air Force first proposed allowing the use of force against aircraft "manifestly hostile in intent" in 1953, leading to adoption of the principle of anticipatory self-defense on the basis of hostile intent. The Navy adopted this principle in 1958, allowing anticipatory self-defense when there was "clear and present danger to the security of the U.S. or its forces."[90] Although this provision was adopted in 1958, it was not a significant factor in U.S. Navy operations until the late 1960s, when the Soviet navy began deploying anti-ship cruise missiles in large numbers.

In 1955 the National Security Council proposed and President Eisenhower approved two of the key provisions in U.S. rules of engagement: the doctrine of hot pursuit and the distinction between hot pursuit and punitive reprisals. Under the doctrine of hot pursuit, U.S. forces could pursue

---

[90]Office of the Chief of Naval Operations, "Protective Measures to be Taken in Applying the Right of Self-Preservation in Peacetime," OPNAVINST 03300.8, February 21, 1958; NWP 9, p. 4-4. On the legal basis for anticipatory self-defense, see John R. Henriksen, "International Claims to Anticipatory Self-Defense: A Juridical Analysis," (Masters Thesis, National Law Center, George Washington University, Washington, DC, 1981), pp. 2-24.

a hostile force out of the area in which they were allowed
to intercept in order to prevent that force from posing a
further threat. Under certain circumstances, hot pursuit
could even be carried into the airspace of another country.
Punitive reprisals, on the other hand, could only be
authorized by the President. Reprisals include attacks
against the territory of the country whose forces had
attacked U.S. forces, and attacks against forces of that
country that were not directly involved in the attack and
which were not an immediate threat to U.S. forces.[91] That
these important concepts originated in the National Security
Council illustrates that civilian authorities have long had
a direct role in shaping rules of engagement.

The next significant changes in the rules of engagement
took place during the Vietnam War. The rules of engagement
became detailed, complex, and cumbersome, requiring positive
identification of targets on the basis of features that were
exceedingly difficult to discern in combat. The rules of
engagement for the air war over North Vietnam were viewed by
military commanders as seriously and unnecessarily
endangering the lives of American pilots. Military
dissatisfaction with what was widely viewed as civilian
"micro-management" of the war led to rules of engagement

---

[91]On hot pursuit and reprisals, see Roach, pp. 50-1;
O'Connell, p. 176; Gallery, p. 25; Thomas C. Schelling, Arms
and Influence (New Haven, CT: Yale University Press, 1966),
pp. 168-70.

gaining a bad reputation among some officers who fought in Vietnam.[92]

During the 1960s and 1970s the scope and complexity of U.S. rules of engagement grew as more sophisticated weapons and electronic systems entered the U.S. and Soviet inventories. The complexity of rules of engagement reflect the complexity of the warfare environment--the more ways in which an adversary can threaten one's forces and the more ways in which one's forces can counter those threats, the greater the number of contingencies that have to be covered by rules of engagement. There were two major reviews of U.S. rules of engagement in the 1970s. The first was in 1973-1975, and entailed deletion of the cumbersome and confusing Vietnam War provisions from the rules as the U.S. withdrew from the war. The second was in 1979-1981, and entailed standardization of the format of the rules among the major commands and expansion of the tactical options for dealing with threats. The 1979-1981 review produced the most significant changes to U.S. rules of engagement since the system of rules originated in the early 1950s. The result was reformulation of the rules of engagement to include a wide range of tactical options that allow rapid

---

[92]For a critique of Vietnam rules of engagement, see W. Hays Parks, "Conventional Aerial Bombing and the Law of War," U.S. Naval Institute Proceedings 108 (May 1982): 98-117; Admiral U.S.G. Sharp, Strategy for Defeat: Vietnam in Retrospect (San Rafael, CA: Presidio Press, 1978); O'Connell, pp. 176-7; Momyer, pp. 133-5, 176-7, 338-9.

and precise tailoring of the rules to meet the political objectives of U.S. leaders while allowing on-scene commanders freedom of action.[93]

Another major review of U.S. rules of engagement was conducted during the mid-1980s, partially in response to the Long Commission's finding that poorly written rules of engagement had contributed to the 1983 disaster at Beirut International Airport. In this review serious attention was devoted to ways in which the rules could better meet the needs of U.S. leaders without creating excessive risks for U.S. forces at the scene of a crisis.[94] No conceptual breakthroughs were made in that review, but the effort appears to have been worthwhile for educating civilian and military authorities as to each other's needs.

Peacetime Naval Rules of Engagement

United States Navy ships and aircraft have an inherent right of self-defense under international law--they may use their weapons to defend themselves if a hostile act of violence is committed against them.[95] The commanding

---

[93]Pease, p. 83; Roach, p. 51; W. Hays Parks, "Crossing the Line," U.S. Naval Institute _Proceedings_ 112 (November 1986): 43.

[94]Cockell interview; Beirut Commission, pp. 44-51.

[95]See "DoD Law of War Program," (DoD Directive 5100.77); _United States Navy Regulations, 1973_, p. 38; and _Commander's Guide_ (NWP 9). Also see Bunn, 69; Burdick H. Brittin, _International Law for Seagoing Officers_, Fifth

officer of a U.S. Navy ship cannot claim that the rules of engagement prohibited him from taking defensive action, for he always has the right and obligation to defend his ship against attack.[96]

The right of self-defense and the conditions under which a commanding officer may use force are defined in United States Navy Regulations. The 1948 edition stated the following:

> 1. The use of force by United States naval personnel against a friendly foreign state, or against anyone within the territory thereof, is illegal.
> 2. The right of self-preservation, however, is a right which belongs to states as well as to individuals, and in the case of states it includes the protection of the state, its honor, and its possessions, and the lives and property of its citizens against arbitrary violence, actual or impending, whereby the state or its citizens may suffer irreparable injury. The conditions calling for the application of the right of self-preservation cannot be defined beforehand, but must be left to the sound judgement or responsible officers, who are to perform their duties in this respect with all possible care and forbearance. In no case shall force be exercised in time of peace otherwise than as an application of the right of self-preservation as above defined. It must be used only as a last resort, and then only to the extent which is absolutely necessary to accomplish the end required. It can never be exercised with a view to inflicting punishment for acts already committed.[97]

The term "friendly foreign state" in the first paragraph was interpreted as meaning any country with which the United

---

Edition (Annapolis, Md.: Naval Institute Press, 1986), pp. 143-144; O'Connell, pp. 53-55, 175.

[96]Roach, p. 49.

[97]United States Navy Regulations, 1948, p. 73.

States was not in a declared state of war. Note that this article allows use of force to protect American citizens and their property when "irreparable injury" is threatened. This was the guidance (amplified by applicable rules of engagement) governing the use of force by U.S. Navy commanding officers from 1948 to 1973.

The 1973 edition of Navy Regulations revised the wording of this article, but left its intent unchanged:

> 1. The use of force in time of peace by United States naval personnel against another nation or against anyone within the territories thereof is illegal except as an act of self-defense. The right of self-defense may arise in order to counter the use of force or an immediate threat of the use of force.
> 2. The conditions calling for the application of the right of self-defense cannot be precisely defined beforehand, but must be left to the sound judgement of naval personnel who are to perform their duties in this respect with all possible care and forbearance. the right of self-defense must be exercised only as a last resort, and then only to the extent which is absolutely necessary to accomplish the end required.
> 3. Force must never be used with a view to inflicting punishment for acts already committed.[98]

The primary difference between the two editions is that the 1973 edition adds the principle of anticipatory defense, allowing use of force to counter an "immediate threat."

Emphasis in Navy Regulations is on caution and restraint. Rules of engagement typically take a similar tone, warning that decisions on the use of force must be "tempered with judgement and discretion." The right of self-defense is recognized in all Department of Defense

---

[98] United States Navy Regulations, 1973, p. 38.

directives related to the law of war and rules of engagement. For example, as Roach points out, rules of engagement always contain a statement to the effect that "Nothing in these rules shall be construed as precluding a commander from using all means at his disposal to exercise the inherent right and responsibility to conduct operations for self-defense of his forces."[99]

Two terms used in rules of engagement are important for understanding the application of the right of self-defense in practice: hostile act and hostile intent. A hostile act is actual use of force--employment of weapons--against a Navy ship. Use of force is always authorized for self-defense when a hostile act is committed. Hostile intent is clear indication that a hostile act is imminent. Demonstration of hostile intent activates the principle of anticipatory self-defense, which authorizes first use of weapons for self-defense when attack is clearly imminent.[100]

The 1981 U.S. Navy confrontation with Libyan forces in the Gulf of Sidra illustrate these rules of engagement provisions in action. On August 18, 1981, a Sixth Fleet battle force built around the carriers USS Forrestal (CV 59) and USS Nimitz (CVN 68) commenced a freedom of navigation exercise in the Gulf of Sidra, on orders from the President,

---

[99]Roach, p. 49. Also see Parks, "Crossing the Line," p. 43.

[100]Bunn, pp. 73-75; O'Connell, pp. 70-71; Roach, p. 50.

to demonstrate United States rejection of Libyan claims of scvereignty over the gulf. The battle force commander, Rear Admiral James E. Service, was dispatched to Washington to brief the JCS and the National Security Council on Navy plans for the operation, including the rules of engagement.[101] The Commander of the Sixth Fleet at the time, Vice Admiral William H. Rowden, described the rules of engagement for the operation: "These rules provided for the right of self-defense; specifically, if fired upon, we had the right, indeed the obligation, to meet force with force. Execution of these rules provided that if we were to fire at any enemy target in self-defense, we intended to hit that target."[102] It was with these rules that the battle force entered the Gulf of Sidra.

Libyan aircraft flew more than 130 sorties against the U.S. ships the first day, but none of the Libyan planes were engaged because they did not fire any weapons against U.S. forces. The next day, however, during an intercept of two Libyan Su-22 Fitters by two U.S. Navy F-14s, one of the Libyan planes fired an air-to-air missile at the Navy jets. This action was met the definition of "hostile act," and

---

[101]Dennis R. Neutze, "The Gulf of Sidra Incident: A Legal Perspective," U.S. Naval Institute Proceedings 108 (January 1982): 28-30; Rowden, p. 270; Hayward, p. 260; Parks, "Crossing the Line," pp. 43-4.

[102]Rowden, p. 271.

the F-14s shot down both Libyan planes.[103] Admiral Rowden

has pointed out an important lesson from this episode:

> At the same time the shoot-down occurred, two other
> intercepts were in progress. All flight leaders were
> able to monitor the engagement on their radios, but no
> one else sought to engage the Libyans with whom they
> were in contact because, beyond the local F-14/Fitter
> incident, there had been no provocation. The rules of
> engagement called only for engagement in self-defense,
> where firing had actually occurred. This Libyan
> incident demonstrates the superb discipline of our
> naval aviators, even when the adrenaline is flowing in
> a crisis situation, and also refutes the notion that we
> are "trigger-happy gunmen" on the lookout to start an
> incident.[104]

Similar restraint was shown by the battle force commander:

> At the time of this attack, ten other Libyan
> fighters were in or near the exercise area. Each gave
> clear indications of hostile intent, as did a Libyan
> Osa missile patrol boat within the exercise area.
> Although his rules of engagement authorized him to
> fire, the task force commander elected not to do so.
> Just as damage to or loss of any part of the task force
> would be translated into a Libyan victory, it is likely
> that Gadhafi would have turned the loss of a missile
> patrol boat and a dozen fighters into a "victory" by a
> martyred David against a bullying Goliath. The
> commander's response was proportionate to the immediate
> threat. His judicious application of force suggests
> the nature of rules of engagement implementation in
> peacetime: however carefully articulated, and
> notwithstanding international legal rights of self-
> defense and a clear designation of authority, there is
> no substitute for the training, experience, and
> judgement of the on-scene commander.[105]

The 1981 episode thus illustrates the manner in which rules

of engagement govern U.S. Navy operations in highly tense

---

[103]Neutze, pp. 26-7; Parks, "Crossing the Line," p. 44.

[104]Rowden, pp. 271-2.

[105]Parks, "Crossing the Line," p. 43.

and volatile situations. When U.S. units needed to use force in self-defense, they had authority to do so. The on-scene commander adhered to the spirit of the rules, which is to use the minimum amount of force necessary for self-defense, rather than to the letter of the rules, which authorized much greater force than was actually employed.

Rules of engagement do not always function as effectively as they did in the Gulf of Sidra in 1981. Norman Friedman and other analysts have noted that there is an inherent tension between excessively tight rules that invite military disaster and excessively loose rules that allow excessively aggressive behavior.[106] Expanding on that idea, rules of engagement can fail in one of two modes: vulnerability failure or escalatory failure. A vulnerability failure is caused by rules of engagement that are excessively restrictive, ambiguous, or complex and confusing. In a vulnerability failure the on-scene commander is unable to take effective action in self-defense, resulting in a successful attacks on his forces that otherwise could have been defeated. An escalatory

---

[106]Friedman, pp. 23-24. Also see Davis, et al., p. 46, who describe the two failures as "military disaster brought about by excessively tight rules, and political catastrophe caused by excessive looseness in the rules." There are three problems with this approach: the military disaster type of failure can have serious adverse political consequences, the political catastrophe type of failure can result in serious military losses, and both types of failure can be caused by excessive ambiguity or complexity in the rules, as well as by excessive tightness or looseness.

failure is caused by rules of engagement that are excessively permissive, ambiguous, complex, or confusing. In an escalatory failure the on-scene commander uses excessive force on grounds of self-defense, causing escalation of the scope or intensity of violence beyond that viewed as desirable by national leaders. Both failure modes have political as well as military consequences, and both can result from rules of engagement that are excessively ambiguous, or complex and confusing.

The deaths of 241 U.S. Marines in a suicide truck bombing of their quarters at Beirut International Airport (BIA) on October 23, 1983 is an example of a rules of engagement failure. It was a vulnerability failure caused by rules of engagement that were excessively ambiguous and restrictive. Responding to guidance from higher authority (originating initially in Washington) that emphasized the high-visibility, non-combat role of the Marines as part of the Multi-National Force, the on-scene commander issued two sets of rules of engagement: permissive rules for Marines guarding the temporary American Embassy (the original embassy had been destroyed by a suicide bomber in September, hence the permissive rules), and restrictive rules for Marines at the airport.

The commission that investigated the disaster concluded that poorly written rules of engagement were a major factor in the disastrous airport bombing. Updating of the rules of

engagement lagged behind the escalation of the threat to the Marines as Lebanese perceptions of the U.S. role in Lebanon shifted from supportive to hostile. Restrictive rules of engagement and the emphasis on their high-visibility, non-combat role created what the Long Commission described as a lax "mind-set" among the Marines at the airport: "In short, the Commission believes the Marines at BIA were conditioned by their rules of engagement to respond less aggressively to unusual vehicular or pedestrian activity at their perimeter than were those Marines posted at the Embassy locations."[107] Consequently, the Marines at the airport were unprepared to counter the suicide truck bomb attack that destroyed their quarters.

In addition to the two rules of engagement provisions described above--a hostile act activating the right of self-defense and hostile intent activating the right of anticipatory self-defense--there is a third provision somewhat broader in scope. Certain designated operational commanders have the authority to declare a force hostile when it presents a "continuing threat of use of force."[108] When a force is declared hostile it can be attacked without

---

[107]Beirut Commission, pp. 50-51. Nearly three decades earlier, Colonel Hadd, initial command r of the Marine landing force in Lebanon in 1958, had pointed out the crititcal importance of staying abreast of a rapidly changing local political environment. See Hadd, p. 86.

[108]Roach, p. 50. Also see Parks, "Crossing the Line," p. 43.

further need to determine hostile intent--additional specific hostile acts or instances of hostile intent are not required to take defensive action. The criteria for declaring a force hostile are quite strict and the authority to declare a force hostile is reserved for senior operational commanders. Thus, this is not a provision that can be used to circumvent the intent of the incident-specific rules. Rather, it is a provision that provides additional tactical flexibility in circumstances of immediate, continuing danger to U.S. Navy ships.

In June 1967, the Commander of the Sixth Fleet, Vice Admiral William I. Martin, used his authority to declare a force hostile in response to reports from USS _Liberty_ that she was under attack by unidentified planes and torpedo boats. After ordering the two U.S. carriers in the Mediterranean to launch eight attack aircraft with fighter escort to defend _Liberty_, he sent the following guidance to the carriers:

> 1. IAW [In accordance with] CINCUSNAVEUR [Commander in Chief U.S. Naval Forces Europe] INST [Instruction] P03120.5B forces attacking Liberty are declared hostile.
>
> 2. You are authorized to use force including destruction as necessary to control the situation. Do not use more force than required. Do not pursue any unit toward land for reprisal purposes. Purpose of counterattack is to protect Liberty only.
>
> 3. Brief all pilots [on the] contents [of] this msg [message].
>
> 4. In addition brief pilots that Egyptian territorial limit [is] only 12 miles and Liberty [is] right on

edge. Do not fly between Liberty and shoreline except as required to carry out provisions [of] para [paragraph] 2 above. Brief fighter cover that any attacks on attack aircraft, Liberty, or they themselves is hostile act and para [paragraph] two above applies.[109]

In a separate message the Sixth Fleet Commander emphasized "Ensure pilots do not repeat do not fly over land."[110]

This episode illustrates three points. First, it shows a fleet commander exercising his authority to declare an unknown force attacking a U.S. Navy ship to be hostile. The pilots sent to defend Liberty were not required to make further judgements concerning hostile intent or the identity of the attackers. Second, it shows a commander exercising prudence by imposing limits on the planes--to not fly over land and to avoid Egyptian airspace--in order to avoid incidents that could escalate the confrontation and have political repercussions. Third, it shows a commander upholding the distinction between self-defense and reprisals. The planes could engage any force threatening Liberty, but were not permitted to pursue attackers in retaliation. Thus, in this instance the rules of engagement provided the on-scene commander with sufficient freedom of

---

[109]Commander Sixth Fleet message, COMSIXTHFLT 081339Z JUN 1967, June 8, 1967 (Declassified 1979. Liberty incident file, Operational Archives, Naval Historical Center, Washington, DC).

[110]Commander Sixth Fleet message, COMSIXTHFLT 081336Z JUN 1967, June 8, 1967 (Unclassified. Liberty incident file, Operational Archives, Naval Historical Center, Washington, DC).

action to exercise initiative in an emergency, but imposed constraints designed to prevent escalation of incidents.

U.S. Navy operational units receive training on standing rules of engagement on a routine basis and intense training in special rules of engagement before commencing operations under those rules of engagement. This training is scenario-based, requiring commanding officers and watch officers to demonstrate their ability to interpret and apply the rules of engagement in various situations. Additionally, the training includes exercises in which rules of engagement situations are simulated, using U.S. Navy units to portray hostile forces, in order to provide a more realistic perspective on how threatening situations develop. The training recognizes that not every possible situation calling for a decision on the use of force can be anticipated, the purpose is to develop the skills of tactical decisionmakers at interpreting the rules of engagement in unfamiliar circumstances. Thus, rules of engagement are not an obscure document that must be hurriedly retrieved from a safe and dusted off when threat of attack becomes imminent. Rather, rules of engagement are an ever-present element in Navy tactical training.

The Decision to Use Force

The on-scene commander or a tactical decision-maker controlling a ship's weapons must make two determinations

when making the decision to use force in self-defense. He
must first determine whether or not the use of force is
authorized under the specific circumstances he faces.
Authorization to use force is a function of three factors:
the national identity of the threatening unit, the national
identity of the threatened target, and the existence of
hostile intent. Having determined that use of force is
authorized, he must then determine the type of defensive
response authorized under the circumstances. These four
factors--identity of threatening unit, identity of target,
existence of hostile intent, and appropriate defensive
response--are the fundamental elements of rules of
engagement.

Identity of threatening force. When a threatening
force has committed a hostile act--firing weapons at a U.S.
ship or plane--there is no requirement that the identity cf
the national attacker be established prior to using force in
self-defense. However, the situation is more complex when a
determination of hostile intent must be made. U.S. rules of
engagement do not necessarily treat the forces of all
potentially hostile nations as being equally threatening.
Special rules of engagement, in particular, can specify
additional precautions against the forces of a nation
perceived as posing a threat to U.S. forces, while leaving
the provisions of standing rules of engagement in place for
the forces of other nations, or even requiring additional

measures to avoid incidents with them. A second, and equally important, consideration is to avoid firing on unarmed civilian vessels and aircraft operating in the vicinity of hostilities. For these two reasons rules of engagement often require identification of potentially threatening ships and aircraft prior to use of force in anticipatory self-defense.[111]

Soviet ships and aircraft, for example, routinely approach close to U.S. Navy vessels at sea with little reaction. The intentions and behavior of Soviet ships and aircraft are well-known, and under normal peacetime conditions they are not an immediate threat to U.S. forces. On the other hand, Iranian planes or warships attempting to approach U.S. Navy ships operating in the Persian Gulf are warned to remain clear, and then fired on if they continue to close.[112] This reflects Iranian hostility toward the

---

[111]Davis, et al., p. 47; Friedman, pp. 39-41.

[112]On U.S. rules of engagement in the Persian Gulf, see "U.S. Confirms Naval Incidents in Strait of Hormuz," New York Times, February 29, 1984, p. A7; U.S. Congress, House of Representatives, Committee on Armed Services, "Report on the Staff Investigation into the Iraqi Attack on the USS Stark," June 14, 1987, pp. 4-6; "U.S. Copters Fire on Iranian Ship," Los Angeles Times, September 22, 1987, p. 1; "U.S. Navy Scares Off Iran's Boats in Gulf," San Jose Mercury News, October 4, 1987, p. 1A; "Stark's Captain 'Failed Fundamentally'," San Jose Mercury News, October 16, 1987, p. 5A; "U.S. f-14 Fighter Fired at Iranian Aircraft," Washington Post, August 11, 1987, p. A1; "U.S.-Iran Missile Encounter: Pure Aerial Electronics," Los Angeles Times, August 13, 1987, p. 6; "U.S. Helicopters Sink 3 Iranian Gunboats in Persian Gulf," Washington Post, October 9, 1987, p. A1.

United States naval presence in the Gulf and prior threatening actions by Iranian forces against U.S. Navy units there. Thus, rules of engagement can be tailored to avoid undesirable incidents with the forces of one country while allowing early and effective defensive action against more threatening forces of another country.

Rules of engagement often specify the certainty of identification required before a determination of hostile intent may be made. A rough scheme of certainty of identification, from greatest to least certainty, would be as follows: positive visual identification by flag or markings, communications intercept, visual identification by class of ship or type of aircraft, electronic intercept (radar, etc.), and pattern of behavior (direction of approach, flight path, formation, etc.). Often more than one indicator present, which can increase certainty of identification. These indicators can be supplemented with intelligence on military vessels and planes known to be in the area. Geography can also aid identification by allowing elimination of forces from countries far removed from the area. The most important distinction drawn in rules of engagement is whether or not visual identification is required prior to using force in anticipatory self-defense.

As was pointed out above, one purpose of rules of engagement is to avoid inadvertenly using force against non-combatant civilian ships and aircraft. However, the

effectiveness of rules of engagement for this purpose can decline significantly once military forces are engaged in combat. On July 3, 1988, the U.S. guided missile cruiser USS Vincennes (CG 48) shot down Iran Air Flight 655 over the Strait of Hormuz. At the time of the incident Vincennes and another U.S. ship had been engaged in a gun battle with Iranian small craft that had attacked a Norwegian ship and fired on a U.S. Navy helicopter. Vincennes sent repeated warnings over international radio channels for the plane to identify itself and state its intentions, all of which were missed or ignored by the Iranian airliner as it flew directly at the U.S. warship. Vincennes misidentified the plane as an Iranian Air Force F-14 jet fighter and the ship shot it down with two surface-to-air missiles. Admiral William J. Crowe, Jr., and other senior naval officers state that the Commanding Officer of Vincennes fully complied with the rules of engagement issued to U.S. forces in the Persian Gulf.[113] The lesson of this incident is that in tense situations--the heat of battle--incidents involving civilian vessels or aircraft can occur even when military commanders

---

[113]"U.S. Downs Iran Airliner Mistaken for F-14," New York Times, July 4, 1988, p. 1; "Statement by Joint Chiefs Head," New York Times, July 4, 1988, p. 4; "U.S. Pushes Inquiry on Downing of Jet," New York Times, July 5, 1988, p. A1; "Senators Assert Warship Captain Reacted Properly," New York Times, July 7, 1988, p. A1; "Navy Won't Alter Engagement Rules," New York Times, July 8, 1988, p. A6; "Errors by a Tense U.S. Crew Led to Downing of Iran Jet, Inquiry is Reported to Find," New York Times, July 11, 1988, p. 1. Also see NWP 9, p. 8-4.

are acting cautiously under rules of engagement designed to prevent such incidents.

Generally, the higher the level of identification that is required, the more difficult it is for Navy ships to take timely defensive measures when threatened. In a high contact-density environment, when the need for positive identification to avoid unwanted incidents is greatest, the identification problem is exacerbated, increasing the danger to Navy ships and the possibility that neutral military forces or civilian vessels or aircraft might be engaged.

Identity of threatened unit. Rules of engagement vary depending upon the identity of the vessel or aircraft being attacked or threatened with attack. Navy units are always allowed to defend against threats to U.S. military forces and U.S. territory, and under most circumstances can defend U.S. civilian ships. Allied military forces can be defended under circumstances defined by arrangements worked out under defense treaties. Rules of engagement covering other types of threatened units--such as civilian ships belonging to allies or the military forces of non-allied friendly nations--are usually quite restrictive. For example, from 1980 to mid-1988 U.S. Navy warships in the Persian Gulf were only permitted to defend U.S. flag merchant ships.[114]

---

[114] Richard W. Murphy, "International Shipping and the Iran-Iraq War," Current Policy No. 958 (Washington, DC: U.S. Department of State, Bureau of Public Affairs, May 19, 1987); Michael H. Armacost, "U.S. Policy in the Persian Gulf

As before, problems can arise in a high contact-density environment where military and civilian vessels in several of the categories are operating in close proximity. Problems can also arise in a rapidly changing political-military environment, when a belligerent suddenly changes the scope of targets he is attacking and rules of engagement lag behind the expanded threat.

Existence of hostile intent. As noted before, a distinction is drawn between hostile acts and hostile intent. Rough categories of indicators of hostile intent, from most to least certainty, are (a) weapon employment (missile or torpedo launch, dropping of bombs, firing of guns--all of which could also be a hostile act in some circumstances), (b) targeting (detection of fire control radar, missile guidance radar, or laser target designation), (c) communications (detection of orders to attack, attack coordination signals, or progress reports), (d) failure to respond to warnings or to comply with declared exclusion zones or broadcast avoidance procedures, (e) exhibiting behavior indicative of imminent weapons employment when in a position to be a threat. Rules of engagement specify, in

---

and Kuwaiti Reflagging," Current Policy No. 978 (Washington, DC: U.S. Department of State, Bureau of Public Affairs, June 16, 1987); "U.S. Denies Scope of Gulf Escort Duty Will Grow," Los Angeles Times, July 9, 1987, p. 8; "Broader Gulf Role is Sought," San Jose Mercury News, October 13, 1987, p. 1A; "Pressure is on U.S. to Act in Gulf," San Jose Mercury News, October 18, 1987, p. 1A.

terms of categories such as these, the level of certainty that hostile intent has been shown that is required before actions can be taken in anticipatory self-defense.

Problems can arise from ambiguous indicators of hostile intent. Hostile forces conducting an actual attack, rather than harassment or a show of force, can be expected to use deception in order to achieve surprise--a highly desirable military tactic that increases the effectiveness of an attack while reducing its costs.[115] Deception tactics include simulating the behavior of non-hostile aircraft and ships, such as by staying within air lanes or sea lanes while in transit, and responding to radio challenges with a civilian identity. Such tactics are particularly effective in a high contact-density environment, and create very difficult rules of engagement problems for tactical decision-makers.

An additional, and even greater, problem is that many indicators of hostile intent can generated by routine peacetime evolutions as well as by actual hostile actions. Spurious indicators of hostile intent can arise during training exercises, weapons testing, and combat systems maintenance. Certain communications and electronic emissions during routine surveillance can also resemble targeting and attack indicators. For example, peacetime

---

[115]Friedman, pp. 39-41.

exercises frequently include tracking and targeting of simulated enemy forces, launching simulated attacks against those forces, and firing of training weapons (which normally do not have warheads). All these activities are essential for maintaining a high level of operational readiness, and by definition generate indicators of hostile intent toward the simulated enemy. The problem is that such indicators of hostile intent may also be detected by units of another country, who may or may not know about the exercise.

Complicating the problem of ambiguous indicators of hostile intent is the Soviet penchant for conducting simulated attacks on U.S. naval forces--a highly dangerous practice. U.S. naval forces have frequently been targets for simulated attacks by Soviet forces, and in some of these incidents only the professionalism and forbearance of the commanding officers prevented a clash. The Soviets have reduced the number of such incidents since the Incidents at Sea Agreement was signed in 1972, but occasionally still conduct simulated attacks.[116] The worst tactical situation

---

[116]Soviet simulated attacks are discussed in Chapter V. Two examples will illustrate the nature of such incidents. In August 1979 in the Black Sea Soviet aircraft, including Backfire bombers, conducted more than thirty simulated missile attacks against the destroyers USS Caron (DD 970) and USS Farragut (DDG 6). On February 18, 1984, again in the Black Sea, a Soviet jet fighter fired its cannon into the wake of the destroyer USS David R. Ray (DD 971). See "Soviet, in 2 Incidents, Takes U.S. torpedo and Baits Ships," New York Times, August 11, 1979, p. 4; "High Seas Diplomacy continuing," Washington Post, February 18, 1984, p. A1. Such actions can indicate hostile intent.

is when Soviet naval forces are in close proximity to U.S. naval forces that are at the scene of a crisis in which one of the local participants is armed with Soviet weapons. A Soviet simulated attack or other weapons training could be mistaken by the U.S. ships as an impending attack by the third party, or similar actions by the third party could be mistaken for an impending Soviet attack. Such situations can easily arise in such perennial hot spots as the Eastern Mediterranean, Persian Gulf, and Sea of Japan.

Ambiguous indicators of hostile intent raise two types of problems. First, suppression of valid indicators through deception can leave ships vulnerable to attack when they would have been authorized to use force in anticipatory self-defense. Second, detection of valid indicators generated by non-hostile activity can result in force being used when it was not, in fact, needed. Thus, the determination of hostile intent is highly context-dependent: the overall political-military environment must also be considered.

Although problems usually arise with the definition of hostile intent, even the concept of hostile act can be troublesome. Is a hostile act actual weapon impact on a defendable category of target, or does it include any weapon employment, even if the target is not hit? If no vessel was hit, to what degree of certainty must it be ascertained that a defendable category of target was the intended target of a hostile act? In a high contact-density environment it may

not be clear which contact was the intended target.
Additionally, inadvertent attacks can occur in the fog of
war and the heat of battle: accidently launched weapons,
mistaken identity, misunderstood orders, indiscriminate
attacks (launched without an effort to identify the target),
and blind impact (unintended target between launch point and
intended target). When a Navy unit is defending itself,
these problems are not at issue. But when other vessels are
to be defended, errors can occur in both directions: failing
to protect a defendable target or taking military action for
a non-defendable target. Both can have serious
consequences.

Appropriate defensive response. The term self-defense
can encompass a wide range of defensive actions, some of
which may be directed (that is, are mandatory when
threatened) and others of which may be prohibited. Rough
categories of defensive actions, in ascending order of
seriousness, are (a) evasion, such as opening the range to
the threat, (b) identification and warning signals to the
threatening unit, informing it that a U.S. Navy ship is
being approached and warning that force may be used in self-
defense, (c) passive measures, such as activating electronic
defense systems (radar jamming) and illuminating the
threatening unit with fire control radars, done as a warning
as well as for their combat value, (d) interposition, such
as placing a U.S. Navy ship between a civilian vessel being

defended and a hostile ship attempting to board or seize it,

(e) warning shots, (f) employment of weapons to destroy the

threatening unit, and (g) pursuit of units that have

committed a hostile act to prevent them from conducting

further attacks. Measures (b) through (e) are not always

required, but may be specified depending on the level of

danger to U.S. Navy units and the likelihood of undesirable

incidents. Even when specified in the rules of engagement,

employment of measures (b) through (e) can be a function of

the time available to execute them before the danger of

attack is acute. This time constraint is recognized in the

rules of engagement, which do not bar the use of force in

self-defense when there is insufficient time to send

warnings or take other passive measures.

Two categories of military action are not authorized

under peacetime rules of engagement: employment of nuclear

or chemical weapons, and retaliatory attacks. Authority to

order the use of nuclear and chemical weapons rests with the

President, and is not pre-delegated in peacetime.[117]

Department of Defense and JCS policy prohibit rules of

engagement from usurping the authority and prerogatives of

---

[117]Roach, pp. 47-8. The one special case is that <u>after</u>
a valid nuclear release order from the President had
authorized employment of nuclear air defense weapons, U.S.
Air Force interceptors carrying nuclear air-to-air missiles
(the Genie rocket, no longer in service) were governed by
special JCS and North American Air Defense Command (NORAD)
rules of engagement for the use of those weapons.

the President in any way. The decision to launch retaliatory attacks for attacks on U.S. Navy units in peacetime is also a matter of national policy, not a tactical decision covered by rules of engagement. If pursuit of attacking units is authorized under the rules of engagement, it is because they represent a further immediate threat to U.S., forces. Pursuit may not be used as a pretext for retaliatory attacks or to justify such attacks afterward.[118]

The Political-Military Context

An understanding the overall political-military context of an incident is essential for the on-scene commander to be able to make decisions on the use of force that support national policy as well as uphold the right of self-defense. In effect, the on-scene commander must determine whether or not the military action he is contemplating will support the political and military objectives and intentions of the President. This requirement is inherent in the two-stage decision process of the rules of engagement system. The political-military context of an incident includes geographic considerations, the political environment, and the overall military situation. These are the same variables that affected the first stage of the decision

---

[118]Ibid, pp. 50-1. Also see NWP 9, p. 6-3; O'Connell, p. 176; Gallery, p. 25; Schelling, pp. 168-70.

process--the original formulation and approval of the rules
of engagement by the President and his military advisors.
These variables influence the political-military intent of
the rules of engagement, which could vary from strenuous
efforts to avoid a military incident at almost any cost, to
a hair-trigger readiness to deliver a sharp response to the
slightest military provocation. The burden on the on-scene
commander is to interpret the intent of the rules of
engagement in the specific circumstances at hand.

Geographic considerations include the proximity of Navy
units to the scene of a conflict or crisis, the proximity of
hostile territory, and the proximity of the conflict or
crisis to U.S. or allied territory. Rules of engagement
attempt to account for geographic factors by issuing special
rules of engagement when it can be anticipated that a ship
will be operating in a high-threat or politically sensitive
area. But when incidents occur while ships are operating
under standing rules of engagement, or when incidents occur
that do not quite fit the circumstances of special rules of
engagement, the on-scene commander must take geography into
account. A demonstration of hostile intent far removed from
the scene of any conflicts is likely to have much different
motives than one committed in the midst of a crisis. For
example, a ship operating near an announced exercise area
has reasonable grounds for concluding that an action
normally regarded as an indicator of hostile intent, such as

being targeted by a fire control radar, is inadvertent
rather than a warning of imminent attack (though it must be
on guard against an accidental attack). The same indicator
could have much more serious implications in the midst of a
military confrontation.

The political environment includes the overall climate
of relations between the U.S. and the nation whose forces
are the potential threat, the stated objectives of the
potentially threatening nation, and U.S. diplomatic
objectives in the conflict. When the potentially
threatening nation is an ally or client of the Soviet Union,
two additional political factors come into play: the overall
climate of relations between the U.S. and the Soviet Union,
and the level of support the Soviet Union is providing to
the potentially threatening nation--overall and in the
specific conflict at hand. Other crises or on-going
conflicts, whether or not they are related to the incident
at hand, are also an important part of the political
environment. World opinion toward the crisis and the U.S.
role in the crisis can be a factor depending on the apparent
responsiveness of the U.S. government to such concerns.
Similarly, domestic political opinion--particularly the mood
of Congress--can be a factor in on-scene decision-making if
emphasized in background briefings. Normally, however,
world and domestic opinion are among the least significant
political influences in rules of engagement decisions.

The overall military situation includes the defense readiness condition (DEFCON) of U.S. forces, other alerts ordered for specific U.S. forces, movements of U.S. forces outside the scene of the crisis, and military incidents involving U.S. forces in other tension areas. The status of these four factors--overall readiness, specific alerts, movements of forces, and other incidents--in the Soviet Union and other nations involved in the crisis are additional major elements the overall military situation. The tactical situation on-scene is the final element in the overall military situation. The tactical situation on-scene is defined by the local balance of military forces, the apparent combat readiness of potentially hostile forces, and the movement of those forces into position for further attacks. Assessment of the military situation is used to distinguish an isolated hostile act (perhaps inadvertent or unauthorized), best answered by a restrained response to ease tensions, from a deliberate provocation or escalation, requiring a sharp response to deter further attacks.

To illustrate the role of the political-military context in the making of rules of engagement decisions, Soviet reconnaissance planes overflying U.S. Navy ships in the open ocean under normal peacetime conditions are not fired on. On the other hand, a Libyan plane attempting to overfly U.S. Navy ships during the March 1986 Gulf of Sidra incident, or an Iranian plane attempting to do so during the

Persian Gulf clashes in 1987 and 1988, would be shot down.[119] Analysis of the political-military context of an incident is a complex task, but U.S. Navy commanders are routinely called upon to do so--without further reference to higher authority--in making rules of engagement decisions.

Because an accurate and comprehensive understanding of the political-military context of an incident is vital to making rules of engagement decisions that support national policy, difficult problems for U.S. forces on-scene can be generated by not informing them of significant military and diplomatic moves. This problem can be particularly acute when rules of engagement are used as a substitute for strategy, that is, when military forces are deployed to the scene of a crisis without a clear mission, only rules of engagement to govern their behavior.[120]

Decisions to change the state of readiness of military forces or to move military forces as a political signal can alter the threat perception and political-military objectives of the target nation, thus altering the tactical situation on-scene. For example, a higher DEFCON can be set

---

[119] On U.S. rules of engagement toward Iran, see the sources in footnote 107. On U.S. rules of engagement toward Libya in 1986, see Parks, "Crossing the Line," pp. 44-52; Lieutenant Commander Robert E. Stumpf, "Air War with Libya," U.S. Naval Institute Proceedings 112 (August 1986): 42-48.

[120] On the importance of keeping the on-scene commander informed, see Train, pp. 301-4. On not substituting rules of engagement for strategy, see Roach, p. 46; Friedman, p. 30.

and troops deployed to forward bases in order to send a
threatening signal, or a lower DEFCON can be set and the
troops withdrawn to send a conciliatory signal. If the on-
scene commander and the tactical decision-makers under him
are not aware of such military actions, the threat to on-
scene forces may be perceived as either higher or lower than
it actually is. Thus, their interpretation of the rules of
engagement may not support national policy--inadvertently
sending conflicting signals to the target nation.

Similar problems can arise from diplomatic initiatives:
by changing the threat perception and objectives of the
target nation, diplomatic moves alter the local tactical
environment. Secret communications and "back-channel"
negotiations are undoubtedly an essential part of diplomacy
and statecraft, but they need to be accompanied by
appropriate efforts to keep the chain of command informed of
the political-military context within which operational
decisions are made.

The Problem of Misinterpretation

As has already been pointed out, the on-scene commander
must interpret the intent of the rules of engagement in the
specific tactical circumstances at hand. He is aided in
doing this by statements of U.S. objectives included in the
rules of engagement when they are issued, by background
briefings on the crisis, and by the tailoring of special

rules of engagement for his specific operation. However, the rules of engagement can still be misinterpreted, producing decisions on the use of force that do not support national policy. These are not tactical decisions made in deliberate violation of the rules of engagement, which is a separate--and exceedingly rare--category of problem. Misinterpretation of the rules of engagement occurs when a tactical decision-maker has an understanding of their intent that is different from the intent of the higher authorities who drafted the rules. U.S. Navy training goes to great lengths to prevent this from happening, but the possibility of misinterpretation cannot be completely excluded.

Misinterpretation of the rules of engagement can arise from three sources. First, verbal orders intended only to emphasize particular operational details can be misinterpreted as a modification to the rules of engagement. For example, a warning that an unusually high number of non-hostile air contacts can be expected could be construed as a requiring greater than normal caution before engaging threatening aircraft. Second, the operational environment can induce routinized patterns of behavior (tactical bad habits, of which complacency is the most common example) that impinge on rules of engagement. For example, daily non-hostile contact with the aircraft of a belligerent could create a routine in which defensive measures authorized in the rules of engagement are not taken

because they have never been needed in the past. Third, a psychological environment can develop that affects threat assessments so as to produce rules of engagement decisions different than intended by the chain of command. The results can be too passive as easily as too aggressive. The constant tension of operating near hostilities can put nerves on edge and generate an intense desire to "do something" rather than continue to be a passive observer or the target of harassment. On the other hand, a feeling that the Navy ships are not needed in the situation (wasting their time) or are impotent to act can lead to complacency and passivity. Personalities can also have an impact--an overly aggressive or cautious tactical decision-maker may provide his own interpretation of the rules of engagement. The professionalism of U.S. Navy officers and their high state of training generally are sufficient to prevent misinterpretation of the rules of engagement, but the on-scene commander and the chain of command needs to be alert for indications of these problems.

Allies and Rules of Engagement

Operations with allies raise further rules of engagement problems. Within the NATO alliance, each member has its own national rules of engagement, which are the rules in effect in peacetime. Additionally, the NATO military command has a separate set of rules of engagement

that govern all member forces after they are transferred to Allied command upon declaration of a NATO Reinforced Alert.[121] This system raises two problems. First, national peacetime rules of engagement may differ among NATO members. This is because threat perceptions, foreign policy objectives, and domestic political constraints can vary considerably among allies. An unambiguous threat from the Soviet Union would undoubtedly be met with a unified response, but a wide range of lesser threats raise political difficulties. Some observers have concluded that diversity in peacetime rules of engagement can interfere with the ability of forces from different NATO members to respond in a consistent and coordinated manner in a crisis. They point out, for example, that while the rules of engagement of most European NATO members require commission of a hostile act prior to use of force, U.S. and British rules of engagement usually permit force to be used upon demonstration of hostile intent.[122] Such differences could create severe difficulties coordinating NATO forces in a sudden crisis.

The second problem is that intense political consultation would almost inevitably be required prior to declaring a reinforced alert, transferring national forces

---

[121]Davis, et al., p. 48.

[122]Drew Middleton, "NATO Issue: When to Let Its Ships Fire," New York Times, April 2, 1984, p. 5; Davis, et al., pp. 48-9.

to Allied command, and bringing NATO rules of engagement into effect.[123] This, some observers fear, could leave NATO forces without clear rules of engagement guidance as a crisis escalates towards war, and could leave them fatally vulnerable to a Soviet conventional pre-emptive strike at the start of a war.[124] Thus, peacetime national rules of engagement formulated for purposes of avoiding war could in fact provide an additional incentive for a pre-emptive attack in a severe crisis.

Summary

In summary, rules of engagement are an important mechanism of indirect control, and are particularly important in crisis management. Rules of engagement are orders issued to define the circumstances in which the U.S. armed forces are authorized to use their weapons for defense against hostile forces in peacetime, and to specify the scope and level of violence of combat operations in wartime. Peacetime rules of engagement prohibit U.S. military commanders from shooting first in peacetime unless absolutely necessary for self-defense. There are two categories of peacetime rules of engagement: standing rules in effect at all times for the forces they cover, and

---

[123]Packenham, p. 49.

[124]Davis, et al., pp. 48-9; Friedman, pp. 32-3.

special rules issued for particularly sensitive situations,
such operations during an international crisis.

Rules of engagement are necessary for five reasons:
first, commanders in the field or at sea often not have the
means or sufficient time to contact higher authorities in an
emergency; second, the lethality of modern weapons has led
to the principle of anticipatory self-defense, so rules are
needed to define indicators of hostile intent; third, not
all nations that are potentially hostile to the U.S. present
the same level of military threat, so rules are needed to
distinguish among them; fourth, the rules ensure that U.S.
forces adhere to the international legal principles of
proportional and minimum force; and fifth, U.S. forces can
be tasked to defend civilian U.S. vessels and U.S. citizens
ashore overseas, the military forces and civilian vessels of
allies and friendly nations, and the territory of allies or
friendly nations.

United States Navy ships and aircraft, like all U.S.
forces, have an inherent right of self-defense under
international law--they may use their weapons to defend
themselves if a hostile act of violence is committed against
them.   Emphasis in U.S. policy on self-defense is on caution
and restraint, and rules of engagement warn that decisions
on the use of force must be "tempered with judgement and
discretion."   The rules of engagement allow force to be used
in self-defense under three circumstances: first, upon

commission of a hostile act against U.S. forces; second, upon demonstration of hostile intent, defined a‍ clear indication that a hostile act is imminent, which activates the principle of anticipatory self-defense; and third, upon an authorized commander declaring a force hostile due a continuing threat of use of force against his command.

An on-scene commander must make two determinations prior to using force in self-defense. He must first determine whether or not the use of force is authorized, which is a function of three factors: the national identity of the threatening unit, the national identity of the threatened target, and the existence of hostile intent. He must then determine the type of defensive response authorized under the circumstances. Making these determinations can be crucially dependent on the political-military context of the operation, which includes geographic considerations, the political environment, and the overall military situation. The on-scene commander must consider whether or not the military action he is contemplating will support the political and military objectives of the President. This makes it critical that on-scene commanders be kept informed of the overall political military situation.

Rules of engagement can fail in one of two modes: vulnerability failure or escalatory failure. A vulnerability failure is caused by rules of engagement that

are excessively restrictive, ambiguous, or complex and confusing. In a vulnerability failure the on-scene commander is unable to take effective action in self-defense, resulting in a successful attack on his forces--an attack that might otherwise have been defeated. An escalatory failure is caused by rules of engagement that are excessively permissive, ambiguous, or complex and confusing. In an escalatory failure the on-scene commander uses excessive force on grounds of self-defense, causing escalation of the scope or intensity of violence beyond that viewed as desirable by national leaders. Either type of failure can result from an on-scene commander misinterpreting his rules of engagement. Misinterpretation of the rules occurs when a tactical decision-maker has an understanding of their intent that is different from the intent of the higher authorities who drafted the rules.

Conclusion

The first objective of this chapter was to explain how delegation and control are exercised in the United States military command system. The previous three sections examined the principles, methods and mechanisms of command and control. The United States armed forces rely on a flexible combination of direct and delegated control. The methods of control range from positive direct control and direct control by negation at the tight end of the

"tightness of control" spectrum, to monitored delegated control and autonomous delegated control at the loose end. Certain of the methods of control can be used in conjunction, and forces can be rapidly shifted from one method to another as the situation warrants. Commanders can exercise indirect control of subordinates even after having delegated them substantial autonomy. This is done via the mechanisms of indirect control: the alert system, standing orders, mission orders, contingency plans, and rules of engagement.

The second objective of this chapter was to to set the United States military command system in the context of organization and management theories on delegation and control in organizations. The first section of this chapter reviewed those theories. Organization and management studies show that significant delegation of decisionmaking authority is common in large organizations. Delegation of decisionmaking is driven by the limits on decisionmaking, which cause decision-making by top-level officials to deteriorate as the size and complexity of the organization increase. These observations apply particularly well to the military chain of command, which is founded on the principle of delegating control while retaining command. As organization theory predicts, delegation of control in the military command system is primarily due to constraints on the ability of top-level authorities to effectively control tactical operations.

Organization and management studies show that tension between autonomy and control is always present in public and business organizations, particularly those consisting of numerous independent operating units. As before, these findings apply particularly well to the U.S. military. Tension between delegation and control is always present in the military chain of command. Pressures toward centralized control are driven by the complexity of modern warfare, fear of nuclear war, and efforts to exploit the force multiplier effect. Pressures toward decentralized control are driven by severe constraints on the ability of top-level authorities to effectively control tactical operations, and by the advantages gained by granting the on-scene commander flexibility to exercise initiative.

Organization and management studies show that delegation of decisionmaking can range from being highly rule-governed, for standard, repetitive situations, to highly discretionary, for situations that cannot be anticipated. This also applies to military command and control. The methods of exercising control cover a "tightness of control" spectrum ranging from very tight to very loose control. Toward the tight end of the spectrum are positive direct control, and direct control by negation. Toward the loose end of the spectrum are monitored delegated control and autonomous delegated control. The guidance contained in mechanisms of indirect control can also range from being

detailed and specific (tight indirect control) to general and flexible (loose indirect control). In military command and control, as in public administration and business management, ighter forms of control are more appropriate for standard situations that are easily anticipated, while looser forms of control are more appropriate for an environment marked by uncertainty and ambiguity, in which specific decisionmaking situations are difficult to anticipate.

Organization and management studies show that three types of control mechanisms are used in various combinations: hierarchical (rules and procedures), collegial (professionalism), and nonhierarchical (organizational and societal norms and culture). All three methods are used in the military organizations. The mechanisms of indirect control--the alert system, standing orders, mission orders, contingency plans, and rules of engagement--are all hierarchical controls. They relieve higher authorities of the burden of having to closely monitor the details of military operations--a burden that can quickly exceed their information processing and decisionmaking capabilities when large-scale operations are being conducted in a fast-paced political-military environment. Relieved of this burden, top-level authorities are better able to concentrate on monitoring the overall political-strategic situation, formulating and revising their strategy for dealing with the

confrontation, and coordinating the overall execution of military operations so that they support that strategy. Hierarchical controls serve similar functions in public and business organizations.

Collegial and nonhierarchical controls have not been discussed, but are more prominent in military organizations than in any other type of organization. Collegial control is provided by the professionalism of the officer corps, which is highly developed and stressed in the training of officers.[125] Non-hierarchical controls--organizational norms and values--are also widely used in the military. They are most visible in elite military units, such as Army Special Forces and the Marine Corps. Members of these units are indoctrinated that their elite status requires that they meet superior standards of performance--typically discipline, endurance, aggressive-ness, and fighting skill-- unique to their organizations. Similar nonhierarchical controls are used throughout the armed forces to complement and reinforce military professionalism.

Collegial and nonhierarchical controls have a major impact on the effectiveness of delegated command and the

---

[125]See Samuel P. Huntington, The Soldier and the State: The Theory and Politics of Civil-Military Relations (Cambridge, MA: Harvard University Press, 1957), pp. 7-18; Morris Janowitz, The Professional Soldier: A Social and Political Portrait (New York: Free Press, 1960), pp. 3-75; Sam C. Sarkesian, Beyond the Battlefield: The New Military Professionalism (New York: Pergamon Press, 1981), pp. 5-54.

mechanisms of indirect control. On the one hand, controls such as discipline, loyalty, and respect for the chain of command are essential for delegated command and the mechanisms of indirect control to function at all. Similarly, professional experience and judgement can be crucial for correctly interpreting ambiguous orders and carrying out general guidance under rapidly changing circumstances. The ultimate test of professional experience and judgement is knowing when to disregard inappropriate orders in order to take action that better supports the national interest. On the other hand, collegial and nonhierarchical controls can generate commitment to particular operational doctrines or procedures, and resistance to operations custom-designed for crisis management purposes. This is the phenomenon emphasized in the organizational process and bureaucratic politics models. A further weakness of those models, then, is that they do not address the full impact--positive as well as negative--of collegial and nonheirarchical controls.

Studies of public administration and business management repeatedly show that in large organizations comprised of numerous independent operating units, optimum performance is achieved with decentralized decisionmaking combined with appropriate--primarily collegial and nonheirarchical--controls. The issue as to what degree of centralization or decentralization is optimum for military

operations was not directly addressed in this review of the military command system. The strength and weaknesses of the methods of control and mechanisms of indirect control, and the arguments for and against centralization of decisionmaking authority, were discussed, but the focus was on how military command and control function in principle. Many things can go wrong in the stress and confusion of crisis military operations, and there are inherent limits on the ability of any methods or mechanisms of control to ensure that decisions made at one level are those that are most appropriate for the situation at another level. The optimum degree of centralization or decentralization can vary widely depending on the nature of the military operation being conducted and the political-military context of the operation. This can be seen in the flexibility of the U.S. military command system and the broad range of control methods--covering the entire tightness of control spectrum--available in it.

One of the central requirements of crisis management is for national leaders to maintain close control over military operations.[126] This requirement can now be addressed in more specific operational terms. National leaders can exercise close control of military operations in a variety of ways. One approach is to shift from methods at the loose

---

[126]George, "Development of Doctrine and Strategy," p. 8; George, "Crisis Management," p. 227.

end of the tightness of control spectrum--autonomous delegated control monitored delegated control--to methods at the tight end of the spectrum--direct control by negation and positive direct control. This is the approach commonly referred to in the crisis management literature. The image of Secretary of Defense McNamara giving rudder orders over the radio directly to Navy ships on the quarantine line during the Cuban Missile Crisis is often viewed as the model of close control that should be followed.

This style of direct control has its costs, and can even impede effective crisis management. Unless the scope of military operations is very small and simple, direct control can quickly overload information processing and decisionmaking. National leaders typically focus on selected aspects of the operations, which may not be the most important or dangerous evolutions taking place. The need for close control thus needs be weighed against the severe constraints on the ability of national leaders to exercise effective direct control of military operations.

A second approach to maintaining close control of crisis military operations is through the mechanisms of indirect control. This entails shifting the guidance contained in mechanisms of indirect control from being general and flexible (loose indirect control), to being detailed and specific (tight indirect control). Close attention to the rules of engagement is particularly

important in this regard. As was also true with methods of control, excessive tightness in the mechanisms of indirect control can be counterproductive--denying the on-scene commander the flexibility he needs to adapt to rapidly changing circumstances. The optimum tightness of control lies somewhere between absolute control and absolute autonomy. Establishing precisely where the optimum balance between control and delegation lies is one of the inherent tensions in crisis management.

U.S. military command and control procedures allow ample opportunity for stratified interaction to occur in crises. The U.S. armed forces rely on a flexible combination of direct and delegated control that emphasizes delegation of authority and providing on-scene commanders with freedom of action. Monitored delegated control is the method of control preferred by military commanders, and when direct control is necessary, control by negation is preferred over positive control. Primary emphasis is placed on use of mechanisms of indirect control rather than on the exercise of direct control. These preferences are strongest in the Navy, which has a long tradition of operational autonomy and which accords "absolute" authority to commanding officers. Even in crises, when there is a tendency for high-level military commanders as well civilian authorities to centralize control over operations, on-scene commanders are delegated substantial decisionmaking authority.

Given that national leaders usually must delegate a certain amount of discretionary decisionmaking authority to military commanders, including the on-scene commander, then it must be expected that military incidents not ordered by national leaders (or even anticipated by them) will occur. The next chapter will explore the range of such incidents that have occurred in the past, and could well occur during a future crisis. Military incidents can generate tactical level military interactions with the forces of the other side in a crisis--interactions decoupled from efforts by national leaders to manage the crisis.

CHAPTER V

TACTICAL-LEVEL MILITARY INTERACTION

The theory of stratified interaction posits that crisis interaction takes place in three semi-independent interaction sequences: political, strategic, and tactical. Thus far, however, little has been said about exactly what types of interactions can occur at the tactical level. It will be useful to explore the nature of tactical-level interactions prior to commencing the case studies so as to identify the types of military actions that are of interest.

The purposes of this chapter are to define the scope of tactical-level interactions that can occur in a crisis and to define analytical categories of crisis military actions. In the remainder of this introduction to the chapter the military actions that can produce tactical-level interactions will be classified, based on the perspective of political-level decisionmakers, as deliberate military actions or inadvertent military incidents. The next three sections will define and present examples of the three categories of inadvertent military incidents: unanticipated authorized actions, military accidents, and unauthorized actions. The fourth section will examine incidents at sea and the impact of the Soviet-American 1972 Incidents at Sea

Agreement. The concluding section will explain why inadvertent military incidents are relatively rare in crises.

Tactical-level interactions will be categorized from the perspective of political-level decisionmakers. Studies of international crises have observed that national leaders generally seek to maintain close control of crisis military operations. The ideal condition that national leaders want to achieve is for no military actions to occur other than the ones they direct. National leaders thus make an an implicit distinction between military actions they initiated deliberately and those they did not order but which occurred anyway. This distinction is the basis for the two major categories of tactical-level interactions: deliberate military actions and inadvertent military incidents.

Deliberate military actions are ordered by political-level decisionmakers. National leaders either issue a direct command for a specific action to be executed, or anticipate the action would occur as a result of an order given previously. Deliberate military actions can thus occur under delegated as well as direct control, and can be ordered in mechanisms of indirect control as well as directly over real-time communications links.

Inadvertent military incidents are military actions that affect the development of a crisis, but which are not specifically ordered or anticipated by national leaders. Avoiding such incidents has been a central concern in the

study of crisis management. As Phil Williams points out, an
"inadvertent outbreak of violence" could cause national
leaders to lose control of events in a crisis:

> The problem is likely to be particularly acute
> if military forces are in close proximity to those of
> the opponent. Trained specifically for warfare,
> military forces are not an ideal instrument in
> situations demanding an enormous degree of caution and
> restraint. Although it is highly improbable that
> hostilities would be initiated without explicit
> orders, a clash between opposing forces resulting from
> the actions of an over-zealous military commander
> cannot be discounted entirely. Thus, policy-makers
> could find themselves losing control over a crisis
> because of the actions of subordinates. This may be
> even more of a problem when geographical distance is
> added to the organizational distance between those who
> formulate and those who execute policy.[1]

Although the importance that national leaders typically

place on maintaining close control over military operations

has been recognized, the questions of how and why

inadvertent military incidents occur and what effect they

have on crisis management efforts have not been adequately

addressed in the literature on crisis management.

Most professional military officers expect that at

least some things, hopefully minor, invariably will go wrong

---

[1]Phil Williams, Crisis Management (New York: John
Wiley and Sons, 1976), p. 100. Two factors mentioned by
Williams--military forces operating in close proximity to
those of the opponent and at great geographical distance
from policymakers--are a good description of typical naval
operations in crises. The one qualification that must be
put on his analysis is that the on-scene military commander
need not be over-zealous for an armed clash to erupt; even a
cautious commander attempting to act with restraint could
become involved in an outbreak of violence that national
leaders did not desire or anticipate.

during military operations. In planning military operations

commanders attempt to allow for mistakes, accidents, and

other unforeseen circumstances. Such problems constitute

what Clausewitz described as "friction" in war--the myriad

things that tend to interfere with the smooth accomplishment

of military operations.[2] One of the principles of war--

simplicity--is followed by military planners because

friction tends to increase with the size and complexity of

an operation. Friction begins to arise as soon as military

operations are launched--well before the enemy is engaged.

It is reasonable to expect, therefore, that friction--things

going wrong--will occur in crisis military operations as

well as in wartime military operations.

Recent studies of crisis management have shown a

growing awareness that inadvertent military incidents are

inherent in the use of military forces as a political

instrument in crises. Eliot Cohen rebuts crisis management

theory, which emphasizes national leaders maintaining close

control of military forces, with the argument that "even if

one were to accept the crisis management theorist's

premises, the statesman must inevitably fall victim to what

Clausewitz called friction, or what we sometimes call

Murphy's Law--the tendency of things to go wrong, of people

not to get messages or to misunderstand or deliberately

---

[2]See Martin Van Creveld, <u>Command in War</u> (Cambridge,
MA: Harvard University Press, 1985), pp. 12-13.

ignore them, of large organizations to fail in their missions for a host of unforeseeable reasons."[3] Cohen goes on to argue that despite this, accidents, misperceptions, and the like are not likely to cause crises to escalate to wars. Robert McNamara, Secretary of Defense during the Cuban Missile Crisis, has proposed a variant of Murphy's Law applying specifically to use of military force in crises:

> I don't think the Cuban Missile Crisis was unique. The Bay of Pigs, Berlin in '61, Cuba, later events in the Middle East, in Libya, and so on--all exhibit the truth of what I'll call "McNamara's Law," which states: "It is impossible to predict with a high degree of confidence what the effects of the use of military force will be because of the risks of accident, miscalculation, misperception, and inadvertence." In my opinion, this law ought to be inscribed above all the doorways in the White House and the Pentagon, and it is the overwhelming lesson of the Cuban missile crisis.[4]

Cohen and McNamara point out the limitations and dangers of crisis management, but do not provide an understanding of the role of inadvertent military incidents in international crises. There is thus a need to take a closer look at such incidents and the impact they can have on crisis management.

---

[3]Eliot A. Cohen, "Why We Should Stop Studying the Cuban Missile Crisis," The National Interest No. 2 (Winter 1985/6): 8. Interestingly, "Murphy's Law"--which states that if something can go wrong, it will--was originated by a military officer (U.S. Air Force Captain Edward Murphy) in 1949. The law has been widely used in the services for almost forty years to teach maintenance personnel to beware of mistakes.

[4]James G. Blight, Joseph S. Nye, Jr., and David A. Welch, "The Cuban Missile Crisis Revisited," Foreign Affairs 66 (Fall 1987): 186.

There are three categories of inadvertent military incidents: unanticipated authorized actions, military accidents, and unauthorized deliberate actions. The next three sections of this chapter will discuss these three types of incidents. Although attention will be focused on incidents involving U.S. forces, examples of accidents involving the forces of other countries will also be presented. This will be done to illustrate that both sides in a crisis can experience inadvertent incidents involving their military forces.

An additional type of incident--incidents at sea--will be discussed in a separate section because they can be either deliberate or inadvertent. Incidents at sea include various forms of harassment and other dangerous interactions between Soviet and American naval forces. They may be initiated deliberately on direct or standing orders from national leaders (for military reasons or as a political signal), or may occur inadvertently--that is, without having been ordered by national leaders. Inadvertent incidents at sea can fall into any of the three categories of inadvertent military incidents: unanticipated authorized actions, military accidents, and unauthorized deliberate actions.

Unanticipated Authorized Actions

Unanticipated authorized actions are military actions taken by military commanders in compliance with guidance

contained in mechanisms of indirect control, but not
directly ordered or specifically approved by national
leaders. Such actions are taken by on-scene commanders in
response to events or tactical conditions that national
leaders did not anticipate, are not aware of, or do not
understand. Such actions are authorized, in that they are
taken in compliance with guidance contained in one of the
mechanisms of indirect control--the alert system, standing
orders, mission orders, contingency plans, or rules of
engagement. But they are unanticipated, in the sense that
national leaders did not directly order the specific action
or anticipate that the specific action would result from
guidance contained in mechanisms of delegated control.
National leaders can only react to an unanticipated author-
ized action and try to manage its impact on the crisis.

The most common phenomenon appears to be that national
leaders order a military operation without understanding the
full range of specific military actions that military
commanders have authority to take in order to carry out that
operation. Ambiguous orders, operations initiated without
specific military objectives to guide decisionmaking by on-
scene commanders, and open-ended military operations (those
that drag on without a definitive conclusion) are particu-
larly prone to cause unanticipated authorized actions.
Reliance on methods or delegated command and mechanisms of
indirect control, although unavoidable for effective control

of a military organization as large as the U.S. armed forces, is the most important condition giving rise to the possibility of unanticipated authorized actions. But such actions can also occur when tighter methods of control are being exercised. National leaders exercising control by negation could approve a military action (by not vetoing it) without understanding what that action entails. This could also occur when positive direct control is being exercised, though in this case it is more accurate to describe the consequences of the action, rather than the action itself, as being unanticipated.

Misperceptions on the part of on-scene military commanders are another possible cause of unanticipated authorized actions. This could occur when a military commander misperceives the political-military context of his local tactical situation. For example, he might misperceive aggressive enemy military moves as indicating that friendly forces are in imminent danger of attack or even that war had started. Believing that attack is imminent or that war has started, the commander takes military actions that would be authorized if one of these situations did, in fact, exist. The possibility of such misperceptions underscores the danger inherent in simulating attacks on an adversary's forces during a crisis--such as the Soviet Navy conducted against the U.S. Sixth Fleet in the Mediterranean while U.S. forces were at DEFCON 3 in the 1973 Middle East War. In

this instance U.S. Navy commanders in the Mediterranean
presumed that Soviet maneuvering to attack was just an
exercise and did not attack any Soviet ships.[5] Under other
circumstances, however, such forbearance could be much more
difficult for on-scene commanders.

Contingency plans can be a source of unanticipated
authorized actions if national leaders do not fully
understand the operational implications of the plans or do
not have the time or inclination to carefully review the
content of a plan before ordering it executed. Although
United States military contingency plans contain a broad
range of options for the employment of military forces,
civilian policy-makers tend to view most predefined military
options as inappropriate because the options were designed
for a crisis scenario different than the one at hand, or
were defined to meet purely military objectives rather than
the requirements for employment of military forces in a
crisis.[6] As was discussed in detail in Chapter IV, there
are inherent limits on the ability of contingency planning.

---

[5]See the case study of the 1973 Middle East War in
Chapter VII of this study.

[6]See Graham T. Allison, Essence of Decision:
Explaining the Cuban Missile Crisis (Boston: Little, Brown
and Co., 1971), pp. 83-89; Leo Hazelwood, John J. Hayes, and
James R. Brownell, Jr., "Planning for Problems in Crisis
Management," International Studies Quarterly 21 (March
1977): 93; Blue Ribbon Defense Panel, Report to the Presi-
dent and the Secretary of Defense on the Department of
Defense (Washington, D.C.: U.S. Government Printing Office,
1970), p. 53.

In practice, top-level military and civilian officials jointly review and revise contingency plans to meet the needs of the specific crisis at hand prior to executing them. However, the possibility of a contingency plan setting in motion military operations that top-level political leaders had not anticipated cannot be excluded entirely.

The alert system can also be a source of unanticipated authorized actions. The President and his advisors--even the Secretary of Defense--may not be aware of the full range of actions that result from setting a higher level of Defense Condition of Readiness (DEFCON).[7] Further, they may not be informed that a particular action has been initiated until it is too late to halt it or until it has already had an unanticipated effect on the crisis. The best example of this was the May 1960 "unintended" DEFCON 3 alert, which was initiated by an ambiguous message from Secretary of Defense Thomas Gates, then in France with President Eisenhower. The message directed the Chairman of the Joint Chiefs of Staff to "quietly order a high state of command readiness," but did not specify a DEFCON level or give sufficient detail for military commanders to determine exactly what the Secretary

[7] John Steinbruner, "An Assessment of Nuclear Crises," in Franklyn Griffiths and John C. Polanyi, eds., The Dangers of Nuclear War (Toronto: University Of Toronto Press, 1979), p. 228. Also see Garry D. Brewer and Paul Bracken, "Some Missing Pieces of the C^3I Puzzle," Journal of Conflict Resolution 28 (September 1984): 454.

of Defense desired. The result was a much greater level of highly visible military activity than had been desired.[8] Such an incident is probably less likely to occur today, but a wide range of unanticipated authorized actions could still result from a presidential decision to set a higher level of DEFCON.

The most important potential source of unanticipated authorized actions is operational decisions made by tactical level military commanders on the basis of guidance contained in standing orders, mission orders, or the rules of engagement. Even when under direct control by top-level political authorities, operational commanders usually have sufficient authority to take actions that could significantly affect the development of a crisis. Ambiguous or ambivalent orders greatly increase the likelihood of unanticipated authorized actions by leaving the on-scene commander uncertain as to the objectives of his mission, the intentions of national leaders, and the actions he is authorized to take. Movement of forces outside the scene of a crisis into battle positions, employment of weapons in self-defense in accordance with the rules of engagement, and stepped up surveillance of sensitive targets are all actions the President might not anticipate as resulting from his decisions, but which could raise tensions in a crisis.

---

[8]Scott D. Sagan, "Nuclear Alerts and Crisis Management," International Security 9 (Spring 1985): 102-6.

Four examples will illustrate the types of unanticipated authorized actions that can occur. The first example concerns two instances of clashes between United States and Soviet forces during the Korean War. Despite the stringent measures that had been imposed on air operations to prevent incidents with the Soviet Union, two serious engagements between United States Navy aircraft and Soviet aircraft took place during the war. The first was on September 4, 1950, over the Yellow Sea. U.S. Navy jet fighters intercepted two Soviet twin-engined bombers approaching a U.S. carrier task force, and when one of the bombers fired on the fighters it was was shot down. The body of one of the bomber's crewmen was recovered, confirming that the plane had been Soviet. The second was on November 18, 1952, over the Sea of Japan. U.S. Navy jet fighters intercepted seven Soviet jet fighters approaching a U.S. carrier task force, shooting down at least two of them.[9] In both cases the carrier task force commanders were authorized to order an intercept to defend their ships under the governing "measures for self-preservation" (the old Navy term for rules of engagement). Neither incident caused a political confrontation between the United States and the Soviet Union, but both had the

---

[9]Malcolm W. Cagle and Frank A Manson, The Sea War in Korea (Annapolis, MD: Naval Institute Press, 1957), pp. 469-475; James A. Field, Jr., History of United States Naval Operations: Korea (Washington, DC: U.S. Government Printing Office, 1962), pp. 167-9, 440-1.

potential to seriously complicate efforts to keep the Korean War limited.

The second example occurred in July 1953, when six ships under the command of Vice Admiral Walter G. Schindler were dispatched to search for survivors from a U.S. Air Force RB-50 reconnaissance plane shot down by Soviet fighters over the Sea of Japan about thirty miles off the coast of the Soviet Union. Vice Admiral Schindler's orders did not specify how close to the Soviet coast he was allowed to search, so he sent the following message to his superiors stating "I intend to take my ships as close to Russian territory as is necessary to recover the airmen from the crashed aircraft" and that he would "brook no interference" from the Soviets.[10] By the time Vice Admiral Schindler received a response directing him to remain clear of Soviet territorial waters, he had already recovered the only survivor that could be located. As it turned out, the survivor was found in international waters, none of Vice Admiral Schindler's ships entered Soviet territorial waters, and there was no harassment of his force by Soviet ships or planes.[11] However, an action like this could produce a

---

[10]Field, pp. 457-9; Rear Admiral Daniel V. Gallery, The Pueblo Incident (Garden City, NY: Doubleday and Co., 1970), p. 57; Vice Admiral Walter G. Schindler, letter to author, March 19, 1988. Vice Admiral Schindler confirmed that he sent the message described by Rear Admiral Gallery.

[11]Field, p. 459; Gallery, p. 58; Schindler letter.

serious Soviet-American confrontation if handled less with

less skill or undertaken in more tense circumstances.

The third example occurred on July 26, 1954 over the

South China Sea. Two days earlier Chinese fighters had shot

down a British air liner en route from Singapore to Hong

Kong. China apologized for the incident, but harassed U.S.

ships and planes engaged in the search for survivors. In

response, the United States moved a task group built around

carriers USS Hornet (CVA 16) and USS Philippine Sea (CVS 47)

into the Tonkin Gulf in a show of force. On July 26, two

Chinese fighters attempting to attack search planes were

shot down by Navy fighters flying combat air patrol for the

search effort.[12] As in the Korean War incidents described

above, the planes acted in compliance with authorized

"measures for self-preservation."

The fourth example occurred during the 1967 Arab-

Israeli War. Sixth Fleet movements on the first day of the

war--a significant signal to the Soviets of American inten-

tions--resulted from a decision made on-scene not related to

---

[12]"2 MIG's Downed British Airliner Off Red China
Coast, Pilot Says," New York Times, July 24, 1954, p. 1;
"Plane Loss Laid to Peiping; U.S. Carriers Rush to Scene,"
New York Times, July 25, 1954, p. 1; "Peiping, Apologizing,
Calls Downing of Plane Accident," New York Times, July 26,
1954, p. 1; "U.S. Rescue Planes Fired Upon, Down 2 Red China
Craft," New York Times, July 27, 1954, p. 1. For details on
the downing of the Chinese planes, see Admiral Harry D.
Felt, "The Reminiscences or Admiral Harry Donald Felt, U.S.
Navy (Retired)," Volume I (Annapolis, MD: U.S. Naval
Institute, Oral History Program, 1974), pp. 268-9.

crisis management efforts taking place in the White House.[13] In this instance, the decision of the Sixth Fleet Commander to move the carrier strike force did not create any problems in managing the crisis. In other circumstances, however, fleet movements unanticipated by national leaders could complicate crisis management efforts by sending unintended political signals to the adversary.

Although none of these examples of unanticipated authorized actions created crisis management problems for American leaders, similar low-level decisions could be more troublesome under other circumstances.

Military Accidents

Military accidents are actions not ordered or deliberately initiated at any level in the chain of command. Military accidents are troublesome because decision-makers may fail to realize they are unauthorized and perceive them as a deliberate provocation, signal of hostile intent, or escalation. This problem is compounded by modern communications systems, which in theory give national leaders in many countries the capability for detailed control of military operations and the ordering specific tactical actions.[14] Richard N. Lebow has warned

---

[13]Jonathan T. Howe, Multicrises (Cambridge, MA: MIT Press, 1971), p. 148.

[14]Williams, p. 202.

that since any military action could conceivably be the result of orders from national leaders, an adversary may assume that those leaders in fact ordered a given action, whether or not the conclusion is warranted.[15] Thus, virtually any military actions can assume strategic importance if believed to have been conceived and personally supervised by national leaders.

In practice, national leaders and even military commanders attempt to distinguish accidents from deliberate provocations or attacks: U.S. naval commanders did so in the 1967 Israeli attack on the Liberty, and Khrushchev did so during the 1962 Cuban Missile Crisis, when a U.S. U-2 strayed over the Soviet Union.[16] Among the factors that are considered when evaluating whether a particular incident was a provocation or an accident are (a) the international political climate (Did the adversary have political and military motives to make a deliberate provocation or attack?), (b) the overall pattern of military operations at the time of the incident (Was the incident isolated or one of several attacks?), and (c) whether the circumstances of

---

[15]Richard N. Lebow, Between Peace and War: The Nature of International Crises (Baltimore: Johns Hopkins University Press, 1981), p. 287. Also see Captain W.T.T. Pakenham, "The Command and Control of Naval Operations: Principles and Organization," Naval Forces 7 (1/1986): 50.

[16]On the Liberty incident, see the case study in Chapter VIII of this study. On the U-2 incident, see Roger Hilsman, To Move A Nation (New York: Doubleday, 1967), p. 221.

of the incident indicate that it was a deliberate action
(Were appropriate combat tactics used?). However, when
assessment of a military accident must be made in the fog of
a crisis, with possibly incomplete and erroneous information
coming in from the scene and decision makers attempting to
sort out adversary intentions under great stress, the possi-
bility of an accident being misperceived as a deliberate
provocation or attack is heightened.

U.S. and Soviet leaders have used communications with
each other to clarify whether incidents were accidents or
provocations. One tactic is to assume (at least for diplo-
matic purposes) that an isolated incident was an accident,
but warn that further such incidents would be viewed as
deliberate provocations or attacks. Khrushchev may have
used this approach when a U.S. U-2 strayed over the Soviet
Union during the Cuban Missile Crisis, warning President
Kennedy that the incident could have had serious conse-
quences. Both of the superpowers have used the "hot line"
to prevent incidents from becoming confrontations: In the
1967 Middle East War the United States informed the Soviet
Union of its military response to the attack on the Liberty,
and in the 1973 Middle East War the Soviet Union protested
the sinking of a Soviet merchant ship during an Israeli raid
on a Syrian harbor.[17] Communications between the United

---

[17]On the 1973 Middle East War incident, see the case
study in Chapter VII of this study.

States and the Soviet Union, particularly over the hot line, have thus proven valuable for sorting out accidents from provocations (and for preventing provocations from recurring by warning against similar "accidents" in the future). Situations could arise, however, in which national leaders or on-scene military commanders on the side that was the victim of a military accident perceive that they do not have time for communications with the other side before taking a military response to an apparent deliberate attack.

An almost infinite variety of military accidents conceivably could occur during international crises. For descriptive purposes, the various types of military accidents will be grouped into the following categories: aircraft incidents, ship and submarine incidents, peacetime weapons incidents, wartime weapons incidents, and miscellaneous incidents. As will be seen in Chapter VII, military accidents rarely occur in crises. Accidents that occurred under noncrisis peacetime conditions and in limited war situations will therefore be used to illustrate the range of military accidents that could possibly occur during a crisis. The reasons why military accidents are rare in crises will be addressed in the final section of this chapter.

Aircraft incidents are the most common type of military accident and can be subdivided into three categories: airspace violations, midair collisions, and

crashes.[18] U.S. military aircraft occasionally have strayed into unfriendly airspace, provoking reactions ranging from diplomatic protests to use of force to down the planes. The following paragraphs will briefly describe all known incidents of this type.[19] The first such incident after World War II occurred on August 9, 1946, when a U.S. C-47 transport that had strayed over Yugoslavia was forced to land. The crew and passengers of the plane were released two weeks later.

The greatest number of incidents occurred during the 1950s. On June 8, 1951, two U.S. Air Force F-80 jet fighters got lost over Germany and landed in Czechoslovakia. The Czechs returned the planes and their pilots a month later.

---

[18]The examples of airspace violations and midair collisions that will be presented all appear to have been inadvertent. It is recognized, however, that both superpowers have motives for taking these actions deliberately: an airspace violation might be viewed as necessary for an especially high priority reconnaissance mission, and a minor midair collision could be used to send a particularly strong political signal. The crashes that will be discussed also all appear to have been accidents, rather than caused by hostile action. It is conceivable, however, that a deliberate downing of an aircraft could be portrayed as an accident, at least by the Soviet Union.

[19]This list is based on published and unclassified sources, and therefore probably is not comprehensive. Not all airspace violations are reported in the press or protested by the country whose airspace was violated. However, any additional incidents were probably very minor and would not affect the findings of this chapter. Excluded from the list are deliberate U.S. airspace violations, such as the U-2 overflights of the Soviet Union conducted from 1956 to 1960, and inadvertent violations of Chinese airspace during the Vietnam War (which will be discussed under wartime incidents).

A U.S. C-47 was forced to land in Hungary on November 19, 1951. The four-man crew was returned several months later. Two U.S. Navy attack planes were attacked by Czech jet fighters on March 12, 1954, close to or inside Czech airspace. The Navy planes escaped with minor damage. There were three incidents in 1958. On June 8 an unarmed U.S. Army helicopter strayed over East Germany and was forced to make an emergency landing. East Germany seized the nine crewmen and passengers, but released them a month later. On June 27 a U.S. Air Force C-118 cargo plane that strayed over Soviet Armenia was fired on but managed to crash land. The nine crewmen and passengers were released by the Soviets twelve days later. On September 2 a U.S. Air Force C-130 transport with seventeen men aboard strayed over the Soviet Union and was shot down. This incident increased Soviet-American tensions because the United States believed that the Soviets had lured the plane over their territory with false radio navigation beacons, because the Soviets refused to admit that they had shot down the plane (the U.S. recorded conversations of the the Soviet fighter pilots during the attack), and because the Soviets refused to return eleven of the bodies to the United States (raising suspicions that the men were being held prisoner).[20]

---

[20]See U.S. Department of State, Historical Office, Bureau of Public Affairs, American Foreign Policy: Current Documents, 1959 (Washington, DC: U.S. Government Printing Office, June 1963), pp. 994-1000; Office of the Chief of

Incidents involving U.S. planes continued to occur in the 1960s. U.S. U-2 reconnaissance planes inadvertently flew over the Soviet Union twice in 1962: on August 30 and October 27 (at the height of the Cuban Missile Crisis). On May 17, 1963, a U.S. Army helo patrolling the demilitarized zone (DMZ) between North and South Korea strayed into North Korean airspace and was shot down, but managed to make a crash landing. The two-man crew was held prisoner for a year before being released. On January 28, 1964, a U.S. Air Force T-39 jet trainer (similar to a civilian Lear Jet) was shot down by Soviet fighters after straying into East German airspace, killing all three pilots aboard the plane. On February 18, 1968, two U.S. Navy jets got lost in bad weather and flew over North Korean territory, prompting a diplomatic protest.[21]

Although the number of incidents involving U.S. aircraft declined significantly in the 1970s and 1980s, they

Naval Operations (OP-09B91R4), "Soviet Attacks on Western Planes," Memorandum, July 15, 1960 (Unclassified. Operational Archives, Naval Historical Center, Washington, DC); Office of the Chief of Naval Operations, "List of U.S. Planes Involved in International Incidents," Memorandum, n.d. (Unclassified, ca. 1965. Operational Archives, Naval Historical Center, Washington, DC. Cited hereafter as OPNAV, "List of U.S. Planes.").

[21]U.S. Department of State, Historical Office, Bureau of Public Affairs, American Foreign Policy: Current Documents, 1962 (Washington, DC: U.S. Government Printing Office, April 1966), pp. 744-45; "U.S. Concedes Air Violation," New York Times, February 19, 1968, p. 9; OPNAV, "List of U.S. Planes."

continued to occur. On July 14, 1977, an unarmed U.S. Army CH-47 transport helicopter strayed north over the DMZ into North Korean airspace and was shot down. Navy jets were again the culprits on April 4, 1983, when six of them accidently overflew Soviet-occupied Zeleny Island in the Kurile chain north of Japan. The Soviets responded by placing their air defense forces in the Far East on alert, filing a diplomatic protest, and overflying the Aleutians with long-range reconnaissance bombers. On April 20, 1984, a U.S. Army AH-1S helicopter on a routine patrol of the West German border strayed over Czechoslovakia and was fired on by Czech jets, but was able to return safely to West Germany.[22]

In summary, U.S. planes or helos were shot down in three of the fifteen incidents (with a loss of twenty-three lives); planes or helos were either forced to land or crash-landed after being shot down, and their crews seized in six

---

[22]Seymour M. Hersh, The Target Is Destroyed (New York: Random House, 1986), pp. 17-19; "Germans Say U.S. Copter Was Over Czechoslovakia," Washington Post, April 22, 1984, p. A17; "Army Concedes Copter Crossed Czech Frontier," New York Times, April 27, 1984, p. 8; "Pentagon Says Its Copter Violated Czech Airspace," Washington Post, April 27, 1984, p. A32. One other incident illustrates the many ways in which such incidents can occur. On July 28, 1966, the pilot of a U.S. U-2 apparently passed out while flying south over the Caribbean from Shreveport, Louisiana. The U-2 continued southward on autopilot until it ran out of fuel and crashed in Bolivia, after flying over Panama, Columbia, Peru and Brazil. See "U-2 Misses Florida Turn, flies to Panama, Vanishes," New York Times, July 29, 1966, p. 1; "Wreckage in Bolivia is Identified as U-2," New York Times, July 31, 1966, p. 4.

of the incidents; and planes or helos were able to escape unharmed or with minor damage in the remaining six incidents. Twelve of the incidents occurred between 1946 and 1968--the period of Soviet-American cold war tensions. Only three incidents occurred after the advent of detente in the early 1970s. As for geographic distribution, seven of the incidents occurred in the airspace of Eastern European countries, five occurred in Soviet airspace, and three occurred in North Korean airspace. U.S. Air Force planes were involved in eight of the incidents (this includes the two U-2 incidents), U.S. Army helos were involved in four incidents, and U.S. Navy planes were involved in three incidents. Two of the incidents (September 2, 1958, and April 4, 1983) significantly increased Soviet-American tensions, though without provoking crises.

Only one of the incidents occurred during a Soviet-American crisis: the U-2 that strayed over the Soviet Union on October 27, 1962, during the Cuban Missile Crisis. Soviet Premier Khrushchev warned that the incident could have been viewed as prelude to an American attack. These findings suggest that inadvertent violations of hostile airspace by U.S. aircraft are not a particularly great threat to crisis management, although they can exacerbate tensions. As a hypothetical example, if the U.S. Army helicopter that strayed over Czechoslovakia on April 20, 1984, had done so during the 1968 during the Soviet

invasion, tensions on the Czech-West German border could have risen considerably.

Soviet and Warsaw Pact military aircraft have frequently violated NATO airspace in Western Europe since the 1950s. In a sensational case, a Bulgarian Mig-17 photographic reconnaissance jet crashed in northern Italy on January 20, 1962. The pilot requested political asylum, claiming that he had flown to Italy in order to defect; the Bulgarian Government claimed the pilot had gotten lost and attempted to land in Italy; the Italian Defense Ministry suspected the plane was attempting to photograph a nearby NATO medium-range ballistic missile (MRBM) base. On July 14, 1966, an armed Soviet helo entered West German airspace and forced an unarmed U.S. Army helo to land. The Soviet helo was driven off when a West German border patrol fired flares at it. The Commander in Chief of U.S. Forces Europe (USCINCEUR) protested the incident to the Commander of the Group of Soviet Forces Germany. Assistant Secretary of Defense for Public Affairs Michael Burch stated in an April 26, 1984, Pentagon press conference that Warsaw Pact planes often violate the airspace of NATO nations, but that the normal Western response was simply to file diplomatic protests.[23]

---

[23] "Photo MIG Down Near NATO Base," New York Times, January 21, 1962, p. 1; "Military in Italy Are Sure MIG Pilot Was Spy," New York Times, January 23, 1962, p. 7; "Soviet Accused in Copter Crash," New York Times, July 16,

In contrast to Soviet violations of Western European airspace, Soviet aircraft have only occasionally violated U.S. airspace. On March 15, 1963, two Soviet Tu-95 Bear reconnaissance bombers violated U.S. airspace over Alaska. U.S. Air Force fighters were scrambled but did not fire on the planes, and the U.S. Government filed a diplomatic protest of the incident. However, Soviet violations of U.S. airspace have been rare since the 1960s. On April 5, 1983, two Soviet Tu-95 reconnaissance bombers entered U.S. airspace over the Aleutian Islands, but Soviet planes normally avoid U.S. airspace.[24]

On the other hand, U.S. fighters also frequently intercept and escort Soviet reconnaissance aircraft entering the U.S. air defense identification zone (ADIZ), which extends hundreds of miles beyond U.S. airspace. Although Soviet aircraft that only enter the U.S. ADIZ are not committing an airspace violation per se, the United States

---

1966; "Army Concedes Copter Crossed Czech Frontier," New York Times, April 27, 1984, p. 8. Soviet reconnaissance planes also frequently violate Japanese airspace. On December 9, 1987, for the first time since the end of World war II, Japanese interceptors fired warning shots at a Soviet plane that had entered Japanese airspace. The Soviet Union apologized for the incident. See "Japanese Jet Warns Soviet Plane," New York Times, December 10, 1987, p. 6; "Japan Given Apology by Soviet," New York Times, December 11, 1987, p. A18.

[24]U.S. Department of State, Historical Office, Bureau of Public Affairs, American Foreign Policy: Current Documents, 1963 (Washington, DC: U.S. Government Printing Office, 1967), pp. 563-64; Hersh, p. 18.

(like many nations, including the Soviet Union) requires prior notification of aircraft entering its ADIZ--a simple, routine procedure normally accomplished by filing a standard flight plan (commercial aircraft) or ADIZ request (military aircraft). U.S. Air Force officials stated in 1983, for example, that fighters scrambled about 1,750 times a year to identify commercial and private planes that inadvertently entered the U.S. ADIZ without prior notification. Soviet military aircraft, not unexpectedly, do not comply with U.S. ADIZ procedures, hence are routinely intercepted when they enter the U.S. ADIZ.[25] There are no reported instances of Soviet aircraft violating the U.S. ADIZ or airspace during a crisis and exacerbating Soviet-American tensions, but the possibility of this happening cannot be discounted entirely.

Midair collisions are infrequent, but happen often enough to warrant attention. At least three such incidents have occurred between Soviet and U.S. or NATO planes since 1970. On March 31, 1970, a U.S. Navy F-4 jet fighter from USS Franklin D. Roosevelt (CVA 42) "brushed" a Soviet Tu-16 Badger reconnaissance bomber over the Mediterranean Sea, causing minor damage to both. In an almost identical

---

[25]See "U.S. Routinely Turns Back Soviet Planes Testing Defense Zone," Washington Post, September 3, 1983, p. A28; "U.S. Fighters Track Soviet Bombers," Washington Post, September 23, 1985, p. A16; "Soviet Planes Intercepted," Washington Post, September 12, 1986, p. A4; "U.S. Fighter Jets Turn Back Soviet Spy Plane Off Alaska," Washington Post, September 22, 1987, p. A4.

incident in October 1973, a U.S. Navy F-4 jet fighter from USS <u>John F. Kennedy</u> (CVA 67) collided with a Soviet Tu-16 Badger reconnaissance bomber over the Norwegian Sea. On September 13, 1987, a Soviet Su-27 jet fighter struck a Norwegian P-3 patrol plane over the Barents Sea, causing the Norwegian plane to lose an engine. In all three incidents both planes involved in the collision landed safely.[26] Ample opportunities for midair collisions arise because U.S. and Soviet reconnaissance planes and interceptors frequently meet in international airspace. In a crisis, a midair collision could well be viewed as a signal of hostile intent, a coercive threat, or an attempt to intimidate. Even worse, if one of the aircraft involved in a midair collision crashed, the side that lost it might think that its plane had been shot down by the other side.

The United States and the Soviet Union have observed each other's aircraft crash in accidents on several occasions over the past forty years. For example, on May 25, 1968, a Soviet Tu-16 Badger reconnaissance bomber crashed in the Norwegian Sea while conducting low altitude surveillance of the USS <u>Essex</u> (CVS 9) ASW carrier group. U.S. Navy helicopters searched for survivors (there were

---

[26]"Naval and Maritime Events January 1970-June 1970," U.S. Naval Institute <u>Proceedings</u> 97 (May 1971): 66; Vice Admiral Joe P. Moorer, Commander of the <u>Kennedy</u> carrier task group in October 1973, letter to author, April 18, 1988; "Soviet Jet Fighter Brushes Norwegian Patrol Plane," <u>New York Times</u>, September 14, 1987, p. 15.

none) and a U.S. destroyer notified a Soviet destroyer of the crash. A Department of Defense spokesman was quick to state that U.S. ships and planes "in no way interfered with, hampered, or threatened the Soviet aircraft at any time or in any way prior to the crash."[27] It appears that in this case and the other cases both sides recognized that the crashes were accidents--at least no public accusations were made that the other side had caused the crash. In a crisis, however, the side that lost a plane in a crash at sea might think that its plane had been shot down by the other side, exacerbating tensions.

Although this study does not address civil aircraft incidents per se, they could exacerbate tensions in a crisis. Civil aircraft have also strayed over hostile airspace and been forced to land or shot down. A partial list of some of the more sensational civilian aircraft incidents shows that they occur often enough to warrant concern: On April 5, 1948 (before the Berlin Crisis erupted in June), a British airliner crashed in Berlin after a midair collision with a Soviet fighter. On April 29, 1952, Soviet fighters fired on an Air France airliner over East Germany. On July 23, 1954, Chinese fighters shot down a British Cathay Pacific airliner over the Pacific. On July 27, 1955, Bulgarian fighters shot down an Israeli El Al

---

[27]"Soviet Bomber Falls After Pass Near U.S. Carrier," New York Times, May 26, 1968, p. 4.

airliner over Bulgaria. On February 21, 1973, Israeli fight-
ers shot down a Libyan airliner over the Israeli-occupied
Sinai Peninsula. On April 20, 1978, Soviet fighters fired
on a Korean Air Lines airliner over the Kola Peninsula,
forcing it to crash land. Incidents involving civilian
airliners are not considered to be tactical-level military
interaction for the purposes of this study, but they can
have political effects similar to those of inadvertent
military incidents.

Two sensational civil aircraft incidents--each
involving one of the superpowers--have occurred in recent
years. The first incident occurred on September 1, 1983,
when the Soviet Union shot down Korean Airlines Flight 007
after it strayed over the Kamchatka Peninsula and Sakhalin
Island. The Soviets tried to justify shooting down the
South Korean plane by claiming that it had been on an
intelligence mission for the United States, and by claiming
that they had identified it as a U.S. intelligence plane.[28]

The second incident occurred on July 3, 1988, when the
U.S. guided missile cruiser USS Vincennes (CG 49) shot down
Iran Air Flight 655 over the Strait of Hormuz. At the time
of the incident Vincennes and another U.S. ship were engaged
in a gun battle with Iranian small craft that had attacked a

---

[28]See Alexander Dallin, Black Box: KAL 007 and the
Superpowers (Berkeley: University of California Press,
1985), pp. 7-15; Hersh, pp. 131-132, 158, 163-165, 171.

Norwegian ship and fired on a U.S. Navy helo. Vincennes

sent repeated warnings over international radio channels for

the plane to identify itself and state its intentions, all

of which were missed or ignored by the Iranian airliner as

it flew toward the U.S. warship. Vincennes misidentified

the plane as an Iranian Air Force F-14 jet fighter and shot

it down. It is clear from statements made by Admiral

William J. Crowe, Jr., and other senior naval officers that

the Commanding Officer of the Vincennes had complied with

the rules of engagement issued to U.S. forces operating in

the Persian Gulf. A Navy inquiry blamed the incident on an

identification error on the part of radar operators.[29]

The Korean Air Lines Flight 007 and Iran Air Flight

655 incidents occurred under much different tactical

circumstances--the Korean plane was shot down in peacetime;

the Iranian plane was shot down in the midst of a battle.

But for this very reason the two incidents illustrate how

civilian aircraft, by being misidentified as military

aircraft or viewed as on a military mission despite civilian

markings, could inadvertently become the target of military

action and exacerbate a crisis.

---

[29]"U.S. Downs Iran Airliner Mistaken for F-14," New
York Times, July 4, 1988, p. 1; "Statement by Joint Chiefs
Head," New York Times, July 4, 1988, p. 4; "U.S. Pushes
Inquiry on Downing of Jet," New York Times, July 5, 1988, p.
A1; "Senators Assert Warship Captain Reacted Properly," New
York Times, July 7, 1988, p. A1; "Errors by a Tense U.S.
Crew Led to Downing of Iran Jet, Inquiry is Reported to
Find," New York Times, July 11, 1988, p. 1.

A wide variety of incidents involving naval ships and submarines can occur on the high seas. Many, perhaps most, incidents are accidents, but others, particularly incidents between Soviet and American naval vessels, are deliberate. This discussion will be limited to accidents, primarily collisions, involving naval ships and submarines. Deliberate incidents will be discussed later, in the section on _ncidents at sea.[30] There will be some overlap in the two discussions because collisions can be deliberate and in some instances it is not clear if the collisions were deliberate or accidental. Furthermore, accidental collisions can occur during deliberate harassment. The focus of this discussion will be on apparent accidental collisions.

Soviet violations of the nautical rules of the road and near collisions with U.S. naval vessels became a serious problem in 1960. Prior to then the Soviet Navy had been very small and limited its operations to coastal waters and adjacent seas. One of first near collisions occurred on April 26, 1960, when the Soviet intelligence collection ship

---

[30]The term "incidents at sea" is used by the U.S. Navy to cover all incidents, whether accidental or deliberate, involving U.S. and Soviet naval vessels and aircraft. The Navy term is used in this study in order to provide an indication of the types of Soviet behavior to which U.S. Navy commanders object. This is important because, as will be seen in Chapter VII, in past crises U.S. civilian authorities have not shown an awareness of what exactly is going on at sea when they order naval forces to the scene of a crisis, and have not understood the dangers that can arise from deliberate Soviet harassment.

(AGI) <u>Vega</u> nearly collided with the U.S. Navy tug USS <u>Nipmuc</u>

(ATF 157) while monitoring submarine-launched ballistic

missile (SLBM) test launches off the coast of Long Island.

This is a good example of an accidental collision nearly

resulting from deliberate behavior--in this case an

extremely aggressive intelligence collection effort. The

United States responded to this and several similar

incidents over the next twelve years with numerous

diplomatic protests, all of which were rejected by the

Soviet Union.[31]

The first actual collision between Soviet and American

naval vessels occurred on June 24, 1966, when the Soviet AGI

<u>Anemometer</u> collided with the intelligence collection ship

USS <u>Banner</u> (AGER 1) in the Sea of Japan. Almost a year

later the second and third collisions occurred, also in the

Sea of Japan. On May 10, 1967, the destroyer USS <u>Walker</u> (DD

517), which was participating in an anti-submarine warfare

exercise, was struck by a Soviet destroyer that had been

---

[31]U.S. Department of State, Historical Office, Bureau
of Public Affairs, <u>American Foreign Policy: Current Docu-</u>
<u>ments, 1964</u> (Washington, DC: U.S. Government Printing
Office, 1967), pp. 669-74; "U.S. Tells Soviet It Imperils
Ships," <u>New York Times</u>, April 4, 1965, p. 1; "U.S. Protests
Harassment of Ships by Soviets; Rejects Soviet Charges,"
Department of State <u>Bulletin</u> 52 (May 3, 1965): 655-58; Naval
Historical Center, "Post-1 JAN 1960 Incidents," Memorandum,
n.d. (Unclassified, ca. April 1965. Post-World War II
Special Lists File, Operational Archives, Naval Historical
Center, Washington, DC. Cited hereafter as "Post-1 JAN 1960
Incidents."). Incidents not otherwise footnoted were found
in Jack Sweetman, <u>American Naval History: An Illustrated</u>
<u>Chronology</u> (Annapolis, MD: Naval Institute Press, 1984).

harassing the U.S. ships. The collision caused only minor

damage to the two destroyers. The next day, May 11, a

second Soviet destroyer collided with the Walker, again with

only minor damage to both ships. The United States promptly

delivered strongly-worded diplomatic protests after each

incident. The Soviet Union, however, blamed the collisions

on U.S. "provocative maneuvers" in the Sea of Japan.

Although some U.S. officials speculated that the two

collisions had been deliberate, perhaps as a political

signal of Soviet displeasure with U.S. involvement in the

Vietnam War, the Captain of the Walker stated that the

collisions appeared to have been accidental.[32]

The third and final collision that occurred prior to

signing of the Incidents at Sea Agreement in 1972 was

between the U.S. destroyer USS Hanson (DD 832) and the

Soviet tug Diomede in the Korean Strait on May 5, 1971.

This minor collision, which was caused by the Soviet tug

violating the nautical rules of the road, did not have

serious repercussions for Soviet-American relations.

---

[32]"A U.S. Destroyer In Far East Bumped By Soviet War-
ship," New York Times, May 11, 1967, p. 1; "A Soviet Warship
Bumps U.S. Vessel 2d Time in 2 Days," New York Times, May
12, 1967, p. 1; "U.S. Plays Down Sea Harassment by
Russians," New York Times, May 13, 1967, p. 11; "U.S. Says
Incident With Soviet Ships Is a 'Closed' Issue," New York
Times, May 14, 1967, p. 1; "U.S. Skipper Views Bumpings as
Error, But Criticizes Soviet," New York Times, May 18, 1967,
p. 1; U.S. Department of State, Historical Office, Bureau of
Public Affairs, American Foreign Policy: Current Documents,
1967 (Washington, DC: U.S. Government Printing Office,
1969), pp. 456-59.

All three of the collisions that occurred from 1966 to 1971 appear to have been accidents. The cause in each case was Soviet violations of the nautical rules of the road and dangerous maneuvers close to U.S. naval vessels. The <u>Banner</u> and <u>Walker</u> incidents were probably cases of accidental collisions during deliberate harassment. The fact that some U.S. officials perceived the <u>Walker</u> incident as deliberate harassment for purposes of political signalling illustrates the potential political impact of accidents at sea.

Although the number of incidents between U.S. and Soviet naval vessels declined significantly after the Incidents at Sea Agreement was signed in 1972, collisions have not been entirely eliminated. In fact, there have been as many accidental collisions since the agreement was signed as there had been before the agreement. On August 28, 1976, a Soviet Echo II nuclear-powered guided missile submarine (SSGN) collided with the frigate USS <u>Voge</u> (FF 1047) in the Mediterranean Sea. The collision was caused by the Soviet submarine, which had been surfaced and on a parallel course with <u>Voge</u>, suddenly turning into the U.S. ship's port side. <u>Voge</u> suffered serious damage to her propeller and had to be towed into port for repairs; the Soviet submarine was damaged but left the scene under its own power. On November 17, 1983, the Soviet Krivak I-class frigate <u>Razyashchey</u>, which had been shadowing the USS <u>Ranger</u> (CV 61) carrier group, collided with the destroyer USS <u>Fife</u> (DD 991) in the

Arabian Sea. Although the Soviet ship caused the collision by violating the rules of the road, Fife had been maneuvering to prevent Razyashchey from approaching too close to Ranger--a tactic aptly described by the Navy as "shouldering." The collision caused only very minor damage to the two ships. The United States filed an Incidents at Sea Agreement protest over the incident. On March 21, 1984, the carrier USS Kittyhawk (CV 63) collided with a Soviet Victor-class nuclear-powered attack submarine (SSN) in the Sea of Japan. The collision, which occurred at night, apparently was caused by the Soviet submarine surfacing directly ahead of the carrier at short range. Kittyhawk suffered minor damage, but the Soviet submarine was forced to remain surfaced and was escorted back to port on the surface--an indication of serious damage.[33]

---

[33] "Soviet Sub and U.S. Frigate Damaged in Crash," New York Times, August 31, 1976, p. 3; "Soviet Frigate Collides With U.S. Destroyer," Washington Post, November 18, 1983, p. A1; "U.S. and Soviet Ships Collide; Navy Says Accident Is Minor," New York Times, November 18, 1983, p. 4; "Soviet Sub Bumps Into U.S. Carrier," Washington Post, March 22, 1984, p. A1; "Soviet Sub and U.S. Ship Collide," New York Times, March 22, 1984, p. 3. In addition to these three accidental collisions, there have two two instances of deliberate Soviet collisions with U.S. ships since 1972 (described in the section of this chapter on incidents at sea). Soviet naval vessels have also collided with the ships of other Western navies. On November 9, 1970, a Soviet Kotlin-class destroyer that had been harassing the British aircraft carrier Ark Royal and her escorts in the eastern Mediterranean struck the side of the carrier after a dangerous maneuver across the carrier's bow. The Soviet destroyer suffered serious damage and lost two men overboard (they apparently died). See "Soviet and British Warships Collide," New York Times, November 11, 1970, p. 2.

There are two differences between the pattern of
accidental collisions prior to the Incidents at Sea
Agreement and the pattern since the agreement was signed.
First, whereas before the agreement two of three collisions
occurred during deliberate Soviet harassment of U.S. naval
vessels, after the agreement only one of three collisions
(the _Fife_ incident) occurred during Soviet harassment.
Second, whereas there were no collisions between ships and
submarines before the agreement, there were two such colli-
sions after the agreement. Both of the ship-submarine
collisions apparently were caused by poor seamanship on the
part of Soviet submarine captains. The overall trend--based
on a very small number of cases--appears to be greater
Soviet Navy caution in surveillance of U.S. Navy ships (less
dangerous maneuvering at close quarters), but more aggres-
sive Soviet use of submarines ʋr close surveillance
(resulting in more ship-submarine incidents). Additionally,
there is an important continuity: the Incidents at Sea
Agreement has not significantly reduced the frequency of
accidental collisions.[34] Despite the agreement's merits, it
has not been sufficient to prevent accidental collisions.

----

[34]A statistical analysis tells us little because of
the small number of cases. Using the overall period in
which serious threat of collisions existed (1960-1987)
yields the following frequency of collisions: one every four
years prior to the Incidents at Sea Agreement (1960-1972),
and one every five years since the agreement (1972-1987). A
single accidental collision in 1988 would lower the post-
agreement rate to equal the pre-agreement rate.

There have been several unconfirmed reports of under-
sea collisions between U.S. and Soviet submarines: in late
1969 involving USS Gato (SSN 615), on March 31, 1971
involving an unidentified U.S. SSN, in May 1974 involving
USS Pintado (SSN 672), in November 1974 involving USS James
Madison (SSBN 627), and in 1986 USS Augusta (SSN 710).
Allegations of additional undersea submarine collisions have
also been made.[35] Although none of these incidents can be
confirmed, they suggest an addition type of naval incident
that could complicate crisis management efforts. Neither
the international nautical rules of the road nor the Soviet-
American Incidents at Sea Agreement govern the behavior of
submarines while submerged (surfaced submarines are clearly
governed by the rules of the road and must remain clear of
other vessels). Interactions between U.S. and Soviet
submarines while submerged are thus regulated only by
policies established by their respective navies, and by the
caution and prudence of their commanding officers.

Accidental collisions, although infrequent, remain a
concern from a crisis management perspective because they
can increase tensions and be misperceived as deliberate

---

[35]"U.S. and Soviet Submarines Are Said to Have Col-
lided," New York Times, January 2, 1975, p. 16; "Submarines
of U.S. Stage Spy Missions Inside Soviet Waters," New York
Times, May 25, 1975, p. 42; "Collision of U.S. and Soviet
Subs Off Siberia in 1974 Is Recounted," New York Times, July
4, 1975, p. 21; "A False Navy Report Alleged in Sub Crash,"
New York Times, July 6, 1975, p. 1; "U.S., Soviet Subs Appar-
ently Collided," Washington Post, March 2, 1987, p. A12.

provocations. The very fact that accidental collisions are relatively rare would make one occurring during a future crisis automatically suspect, particularly if the ship that is the victim of the collision suffered much greater damage than the ship that caused the collision. If a U.S. destroyer had collided with a Soviet submarine at the height of the 1962 Cuban Missile Crisis (Say, on October 27), severely damaging or sinking the submarine, President Kennedy probably would have had difficulty convincing Khrushchev that the incident was an accident. Khrushchev might well have viewed the collision as deliberate U.S. retaliation for the shooting down of an American U-2 that same day. Such a hypothetical incident might not have prevented resolution of the crisis short of war the next day (October 28), but certainly would not have made that resolution any easier.

Collisions also occur between naval ships or submarines and civilian vessels. On February 1, 1968, the U.S. destroyer USS <u>Rowan</u> (DD 782) collided with the soviet merchant ship <u>Kapitan Vislobokov</u> in the Sea of Japan. At the time, the U.S. Navy was conducting high-tempo operations in the Sea of Japan in the wake of the North Korean seizure of the USS <u>Pueblo</u> (AGER 2). On April 9, 1981, the U.S. submarine USS <u>George Washington</u> (SSBN 598) collided with the Japanese merchant ship <u>Nissho Maru</u> in the East China Sea, sinking the ship and killing two of its crewmen. This was one of at least eight collisions between U.S. submarines and

civilian merchant ships from 1965 to 1982, but the only one in which the merchant ship was sunk. Of the eight collisions, four involved SSBNs, two involved SSNs, and two involved conventional attack submarines (SS). None of the U.S. submarines was seriously damaged. No Soviet merchant ships were involved in any of the collisions with U.S. submarines (the ships were of Norwegian, Lebanese, West German, Philippine, Japanese, Turkish, and U.S. registry).[36]

Submarines belonging to the Soviet Union and other countries have also collided with civilian vessels. On September 21, 1984, a Soviet Victor-class SSN collided with a Soviet merchant ship in the Strait of Gibralter, seriously damaging the submarine and apparently sinking the merchant ship. On at least two occasions Western submarines have collided with Soviet Bloc merchant ships: on September 1, 1976, a Turkish submarine collided with a Soviet freighter, and on January 21, 1983, a West German submarine collided

---

[36]On the _Rowan_ incident, see "Soviet and U.S. Ships Collide," _New York Times_, February 3, 1964, p. 4. On the submarine-ship collisions, see "A-Powered Sub Damaged, Hit By Turkish Ship," _Washington Post_, March 24, 1982, p. A12; "Naval and Maritime Events 1981," U.S. Naval Institute _Proceedings_ 108 (May 1982): 62; "Naval and Maritime Events, 1 July 1968-31 December 1969," U.S. Naval Institute _Proceedings_ 96 (May 1970): 52; "Sub in Collision Off Spain," _New York Times_, August 10, 1968, p. 27; "U.S. Nuclear Submarine and Freighter Collide," _New York Times_, January 11, 1965, p. 41; "U.S. Submarine Collides With Freighter Off China," _New York Times_, August 12, 1965, p. 54. The author was told by a retired senior naval officer that there was a collision between a U.S. submarine and a Soviet merchant ship in 1965, but was unable to confirm the incident.

with an East German passenger liner.[37] Thus, consideration

of maritime accidents that could affect international crises

must include the possibility of collisions between

submarines and merchant ships.

During a crisis, an accident involving a naval vessel

of one side and a civilian merchant ship of the other side

could be viewed as a deliberate provocation in two

situations: First, if the naval vessel sank and the cause

was not immediately known, the side that lost the ship could

well suspect hostile action by the adversary. Second, if

the merchant ship was sunk or seriously damaged, and had

been carrying military supplies or other critical materials

to an ally directly involved in fighting, the side that

owned the ship could well suspect that the collision was a

deliberate attempt to prevent delivery of its cargo.

---

[37]"Soviet Sub, Merchant Ship Collide, Damaging Both in Gibralter Strait," Washington Post, September 22, 1984, p. A16; Jane's Defense Weekly, October 20, 1984, p. 667; "Soviet Ship in Collision," New York Times, September 2, 1976, p. 16; "Ship that Hit Andrea Doria is Hit," New York Times, January 22, 1983, p. 2. There have also been two reported instances of Soviet submarines colliding with Soviet surface combatants on the high seas. See "Damaged Submarine Is Sighted in the Caribbean," New York Times, September 6, 1973, p. 74; "Soviet Sub Damaged in the Mediterranean," New York Times, February 4, 1970, p. 8.

[37]Destroyers and frigates are relatively small vessels that can be seriously damaged by large merchant vessels. For example, on June 6, 1971, the Soviet tanker Busharov collided with the French frigate Surcouf in the Mediterranean. The frigate's bow was sheared off and sank, nine of her crewmen were killed, and what was left of the ship had to be towed into port. See "9 Lost as Soviet Tanker Hits a French Warship," New York Times, June 7, 1971, p. 6.

Peacetime weapons accidents occur primarily during training exercises and maintenance or testing of weapons systems. Weapons accidents can also be a collateral result of other emergencies, such as an aircraft jettisoning ordnance to lower its weight after losing power or a ship jettisoning ordnance threatened by a fire. Jettisoned weapons normally do not cause as much damage as deliberately launched weapons, but can still cause substantial damage to ships or aircraft. Three types of peacetime weapons incidents will be discussed: tactical missile accidents, naval gunfire accidents, and torpedo accidents.

Tactical missiles are particularly accident-prone, and on several occasions have been accidently launched or have gone astray after being deliberately launched. Air-to-air missiles, probably the least dangerous due to their relatively short range, have been involved in several accidents: On May 27, 1974, a U.S. Navy A-7 shot down a Navy A-4 off the coast of Florida; on July 22, 1974, a Navy F-4 shot down a Navy helicopter off the Philippines; and on September 22, 1987, a Navy F-14 accidently shot down an Air Force F-4 during an exercise in the Mediterranean.[38] On at

----

[38]"Naval and Maritime Events, January 1974-June 1974," U.S. Naval Institute Proceedings 101 (May 1975): 61; "Naval and Maritime Events, July 1974-December 1974," U.S. Naval Institute Proceedings 101 (May 1975): 141; "Navy F14 Downs Air Force Jet During Exercise," Washington Post, September 23, 1987, p. A4. The September 22, 1987, incident was caused by an inexperienced pilot mistaking an order to conduct a simulated attack as an order to conduct a real

least two occasions stray air-to-air missiles have struck ships: On August 12, 1968, a Navy missile hit a civilian oil survey vessel off the coast of California, setting it afire and injuring three crewmen; and on July 29, 1986, a civilian tanker was hit by a Navy missile off the coast of Virginia, causing a small fire but no injuries. On another occasion a Navy Bulpup air-to-surface missile accidently hit a Navy rescue craft off the coast of California, sinking the vessel but miraculously not injuring any of the crew.[39]

Anti-ship cruise missiles have also been involved in accidents. On July 14, 1981, the U.S. Navy guided missile destroyer USS Coontz (DDG 40) accidently launched a Harpoon missile in the Caribbean Sea during routine system testing. The missile crashed into the sea at the end of its flight without striking anything. On December 28, 1984, an unarmed Soviet SS-N-3 missile launched during an exercise in the Barents Sea went astray, flew over northern Norway, and

attack. See "A War Game That Turned Real," San Jose Mercury News, April 11, 1988, p. 1A. This is significant because the Soviet Navy frequently conducts simulated attacks on U.S. Navy ships, and did so even at the height of Soviet-American tensions during the 1973 Middle East War (See Chapter VII). During an acute crisis, an anti-ship missile mistakenly launched at a U.S. Navy ship by an inexperienced Soviet pilot could easily be misperceived as a deliberate attack and provoke a sea battle, particularly if other Soviet units were simulating attacks at the same time.

[39]"Naval and Maritime Events, 1 July 1968-31 December 1969," U.S. Naval Institute Proceedings 96 (May 1970): 28; "Navy Missile Damages Tanker Off Norfolk," Washington Post, July 30, 1986, p. A10; "Navy Ship Sunk by Errant Missile," New York Times, February 7, 1968, p. 34.

crashed in Finland. The Soviet Government apologized for

the incident.[40] Accidents with tactical missiles have the

potential to cause an incident because both the United

States and the Soviet Union (and their allies) routinely

monitor the other side's naval exercises with ships and

aircraft. Although greater caution is shown during crises,

the possibility of an incident with a tactical missile

cannot be eliminated.

Strategic nuclear missiles--intercontinental ballistic

missiles (ICBMs), submarine launched ballistic missiles

(SLBMs), and long-range land-attack cruise missiles--are the

safest of all missiles because of their greater safeguards

against accidental launch, but are not immune to accidents.

On May 9, 1973, a U.S. Navy Poseidon SLBM went out of

control during a test flight and crashed in the Atlantic

---

[40]"Naval and Maritime Events, 1982," U.S. Naval
Institute Proceedings 108 (May 1982): 87; "Soviet Cruise
Missile Said to Stray Across Norway and Into Finland," New
York Times, January 3, 1985, p. 1; "Norway and Finland
Report Moscow Apology on Missile," New York Times, January
5, 1985, p. 4. The U.S. Harpoon missile carries only a
conventional warhead; the Soviet SS-N-3 missile can carry
either a nuclear or a conventional warhead. Lest it appear
that such accidents are exclusively a Navy domain, two other
incidents should be mentioned: On January 4, 1967, an
unarmed U.S. Air Force Mace land attack cruise missile being
used as a target drone went astray, flew over Cuba, and
crashed in the Caribbean; and on September 12, 1967, an
unarmed U.S. Army Pershing tactical ballistic missile went
off course during a test flight and crashed in Mexico. See
"U.S. Target Missile Strays Over the Tip of Cuba," New York
Times, January 5, 1967, p. 19; "Unarmed Pershing Missile
Strays Into Mexico," New York Times, September 13, 1967, p.
24. Operational versions of the Mace and Pershing missiles
were armed with nuclear warheads.

near the Soviet AGI *Zakarpatye*. One of the more spectacular strategic missile incidents occurred on September 11, 1986, when a Soviet SS-N-8 SLBM on a routine test flight crashed in a remote area of northeastern China.[41] Although an accident involving an operational strategic nuclear missile could have catastrophic consequences during an international crisis, the probability of such an incident appears to be extremely low due to the safeguards against accidental or unauthorized launch of strategic missiles. Additionally, an accident during a strategic missile test flight probably would not be misperceived as a deliberate attack because the superpowers carefully monitor each other's test flights.

The remaining types of peacetime weapons accidents—naval gunfire accidents and torpedo accidents—rarely have consequences as serious as tactical missile accidents, but could still exacerbate tensions in a crisis. On March 8, 1963, and February 5, 1979, shells fired by U.S. Navy ships during training exercises fell in the vicinity of Soviet vessels that were near or inside publicly announced training areas.[42] A Soviet Navy Foxtrot-class submarine being towed

---

[41]"Naval and Maritime Events, January 1973-June 1973," U.S. Naval Institute *Proceedings* 100 (May 1974): 58; "A Soviet Missile Is Said to Misfire and Hit China," *New York Times*, September 16, 1986, p. 1.

[42]U.S. Department of State, Historical Office, Bureau of Public Affairs, *American Foreign Policy: Current Documents, 1963* (Washington, DC: U.S. Government Printing Office, 1967), pp. 563-63; "Naval and Maritime Events, 1979," U.S. Naval Institute *Proceedings* 100 (May 1980): 50.

to Cuba was involved in the second incident. Torpedo accidents normally occur during exercises and testing: On April 23, 1958, the destroyer USS Yarnall (DD 541) was struck by a torpedo fired during an exercise, and on December 19, 1983, the frigate USS Jack Williams (FFG 24) accidently launched a torpedo that landed on the pier next to the ship but did not explode. A more dangerous incident occurred on October 7, 1973, when a Soviet Kanin-class destroyer that had been shadowing a British aircraft carrier launched a torpedo in the midst of a NATO naval exercise. The Soviet destroyer had experienced an explosion and apparently jettisoned the torpedo to prevent it from being engulfed in the ensuing fire.[43] Although it was clear in this case that the torpedo launch was not a deliberate hostile act, a naval gunfire or torpedo incident at the height of a crisis could well be misperceived as an act of war—particularly if the shells or torpedo struck a warship belonging to the other side.

Once shooting starts, the probability of serious accidents greatly increases. Aircraft can easily stray over national borders into hostile airspace from which they had

---

[43]"Destroyer Docks for Repairs," New York Times, April 24, 1958, p. 27; "U.S. Navy Ship Sights Pier, Torpedoes Same," Washington Post, December 20, 1983, p. A3; "Naval and Maritime Events, July 1973-December 1973," U.S. Naval Institute Proceedings 100 (May 1974): 295; "Soviet Ship Fires Torpedo During NATO Exercise," The Times (London), October 8, 1973, p. 1.

345

been excluded. During the Korean War U.S. planes were forbidden from flying over Chinese or Soviet territory, but occasionally strayed over both countries. On at least two occasions U.S. planes accidently attacked targets in China. The most serious incident occurred on October 8, 1950, when two U.S. Air Force F-80 jet fighters strayed over the border and strafed a Soviet airfield.[44] U.S. planes inadvertently entered Chinese airspace on several occasions during the Vietnam War and China publicly protested hundreds of alleged violations. There were no reported instances of U.S. planes attacking ground targets in China, but at least eight U.S. planes were shot down and one U.S. pilot captured by the Chinese.[45] Incidents similar to those that occurred in the Korean War and Vietnam War are to be expected when-ever sustained air combat operations are conducted close to international boundaries.

Inadvertent or indiscriminate attacks on naval vessels and civilian merchant ships also occur occasionally in

[44]Robert F. Futrell, The United States Air Force In Korea, 1950-1953 (New York: Duell, Sloan and Pearce, 1961), pp. 142, 208-11, 453, 567.

[45]"Chinese Reds Say Fighters Downed U.S. Attack Plane," New York Times, April 13, 1966, p. 1; "China Charges Intrusion By U.S. Ship and Planes," New York Times, September 26, 1966, p. 2; "U.S. Says Plane Flew Over Hainan Accidently," New York Times, February 10, 1967, p. 2; "2 U.S. Navy Jets Downed in China; One Pilot Seized," New York Times, August 22, 1967, p. 1; "Unarmed Navy Plane Downed by a Chinese Communist MIG," New York Times, February 15, 1968, p. 2; "467th Warning by Peking," New York Times, March 28, 1969, p. 40.

limited war situations. Two serious cases of U.S. planes attacking U.S. and allied naval vessels occurred during the Vietnam War. On the nights of June 16-17, 1968, U.S. Air Force fighters attacked radar targets that they believed were low-flying North Vietnamese aircraft or helicopters over South Vietnam. The Air Force planes actually fired on U.S. and Australian Navy ships off the coast of South Vietnam, sinking a U.S. Navy gunboat (PCF 19), and damaging the cruiser USS Boston (CAG 1), the destroyer USS Edson (DD 946), and the Australian destroyer HMAS Hobart (D 39). Seven American and two Australian sailors died in the Air Force attacks.[46] The second incident occurred on April 16, 1972, when a U.S. plane inadvertently fired a Shrike anti-radar missile at the USS Worden (DLG 18), seriously damaging the ships's antennas and superstructure and killing or wounding thirty crew-men.[47] Inadvertent attacks by U.S. forces against other U.S. forces dramatically illustrate the dangers inherent in limited war situations.

---

[46]"Naval and Maritime Events, 1 July 1968-31 December 1969," U.S. Naval Institute Proceedings 96 (May 1970): 24; Australian Department of Defence, Navy in Vietnam (Canberra: Australian Government Printing Service, 1980), pp. 59-60; George Odgers, The Royal Australian Navy: An Illustrated History (Hornsby, New South Wales: Child and Henry, 1982), p. 181; Eugene G. Windchy, Tonkin Gulf (Garden City, NY: Doubleday and Co., 1971), p. 262.

[47]"Naval and Maritime Events, January 1972-June 1972," U.S. Naval Institute Proceedings 99 (May 1973): 56; "'72 Crippling of U.S. Ship Cited," Baltimore Sun, June 5, 1980, p. A4.

Soviet ships have been attacked on at least three occasions during limited war and crisis situations. On June 2, 1967, two U.S. Air Force F-105 fighter-bombers on a raid over North Vietnam accidently attacked the Soviet freighter Turkestan in a North Vietnamese port. Knowing that they had violated strict regulations against attacks on foreign merchant ships in North Vietnamese ports, the pilots responsible for the attack initially attempted to hide their mistake. Consequently, the U.S. Government denied that American planes were to blame after the Soviet Union protested the incident.[48] On October 17, 1969 South Vietnamese Navy gunboats fired on a Soviet intelligence collection ship (AGI) allegedly in South Vietnamese waters near Danang. The Soviet vessel fled the scene trailing smoke.[49] This was a deliberate attack, rather than an accident, but illustrates the type of unanticipated incidents that can occur when superpower naval vessels

---

[48]"Moscow Says U.S. Hit A Soviet Ship in Vietnam Port," New York Times, June 3, 1967, p. 1; "U.S. Denies Attack on Soviet Vessel in North Vietnam," New York Times, June 4, 1967, p. 1; "Moscow Rejects Denial of Attack," New York Times, June 6, 1967, p. 3; "Colonel Destroyed Proof U.S. Strafed Soviet Ship," New York Times, April 11, 1969, p. 11; Phil G. Goulding, Confirm or Deny: Informing the People on National Security (New York: Harper and Row, 1970), pp. 139-52.

[49]"Saigon reports Its Patrol Boats Fired At and Hit Soviet Spy Ship," New York Times, October 18, 1969, p. 1; "Thieu Confirms Attack on Soviet Ship," New York Times, October 19, 1969, p. 3; "Naval and Maritime Events, 1 July 1968-31 December 1969," U.S. Naval Institute Proceedings 96 (May 1970): 545.

operate near a war zone. A third incident occurred during the 1973 Middle East War. Israeli missile boats raided the Syrian port of Tartus the night of October 11-12, sinking two Syrian missile boats and the Soviet freighter _Ilya Mechnikov_ with anti-ship missiles. Israel expressed regret for sinking the Soviet ship and claimed that its forces had orders not to attack civilian vessels. In a message delivered to the U.S. on October 12, the Soviet Union protested the Israeli attack and warned the "The Soviet Union will of course take measures which it will deem necessary to defend its ships and other means of transportation."[50] Inadvertent or indiscriminate attacks on merchant ships or naval vessels could easily exacerbate tensions in a crisis.

The final category of military accidents is miscellaneous accidents. Three types of accidents in this category will be discussed: sinkings of naval vessels other than those caused by collisions or the adversary's weapons, explosions at shore bases, and electromagnetic interference incidents.

---

[50]"Israel Is Accused in U.N. of Sinking a Soviet Ship," _New York Times_, October 13, 1973, p. 1; "3 Freighters Sunk," _New York Times_, October 13, 1973, p. 1; Henry Kissinger, _Years of Upheaval_ (Boston: Little, Brown and Co., 1982), p. 510.

Sinkings of naval vessels can result from causes other than collisions or an adversary's weapons. The most common examples are sinkings of Soviet submarines, which are notoriously accident-prone. For example, on October 7, 1986 a Soviet Yankee-class SSBN on patrol in the Atlantic suffered an explosion and fire in its missile compartment. The submarine was able to surface, but sank three days later while under tow.[51] Soviet surface ships have also been sunk in accidents. On August 19, 1970, a Soviet AGI that had been monitoring a NATO naval exercise in the North Sea capsized. A Soviet tug in the area rescued the crew. In September 1974, the Soviet Kashin-class guided missile destroyer Otvazhny exploded, burned, and sank in the Black Sea, with the loss of almost the entire crew.[52] Incidents like these could cause serious tensions during a crisis if the Soviet ship or submarine sank without survivors and U.S. Navy units in the area were suspected by the Soviets of complicity in the unexplained loss of the vessel.

Major explosions at shore bases could also cause crisis management problems. On May 13, 1984, explosions and

[51]Norman Friedman, "World Naval Developments," U.S. Naval Institute Proceedings 112 (December 1986): 122; Jim Bussert, "The Safety of Soviet Nuclear Submarines," Jane's Defence Weekly, April 18 1987, pp. 715-720.

[52]"Naval and Maritime Events, July 1970-December 1970," U.S. Naval Institute Proceedings 97 (May 1971): 337; "Soviet Ship Capsizes Near Nato exercise," New York Times, August 20, 1970, p. 14; "Soviet Destroyer Is Reported Sunk," New York Times, September 27, 1974, p. 7.

fires destroyed a Soviet naval tactical missile storage site in the Severmorsk base complex on the Kola Peninsula.[53] Had this disaster occurred at the height of the Cuban Missile Crisis, possibly at the same time an American U-2 was being pursued over Siberia, it might have triggered a Soviet perception of an American attack on the Soviet Union. Rational analysis of such an explosion would quickly lead to the conclusion that had not been of U.S. doing: U.S. forces would attack air defense sites and operational forces before ordnance storage sites, and a single, isolated attack would be highly unlikely. But in an acute crisis, with the two sides on the brink of war and military commanders on both sides focusing on worst-case scenarios, rational analysis could succumb to the effects of tension and stress. Furthermore, if it is safe to assume that the probability of an accident at an ordnance storage site is directly proportional to their level of activity, and that preparations for war during an acute crisis include delivery of ammunition to operational forces as their readiness is increased, then the contention can be made that an accident like this is more likely to occur during a crisis than at any other time.

Electromagnetic interference incidents are not uncommon, but normally do not cause serious problems and

---

[53]"Soviets' Northern Fleet Disabled, 'Not Viable' for Six Months," Jane's Defense Weekly, July 14, 1984, pp. 3-4; Donald C. Daniel and Gael D. Tarleton, "The Soviet Navy in 1984," U.S. Naval Institute Proceedings 111 (May 1985): 362.

therefore are rarely reported in the press. Electronic warfare measures, such as jamming of radars or communications, commonly conducted for training or maintenance, have in the past inadvertently degraded other countries' weapons systems. Chaff, a cloud of metal particles used to confuse radars, has on occasion drifted astray, interfering with civilian air traffic control radars and even causing power outages. On at least one occasion Soviet Tu-95 Bear reconnaissance bombers operating off the east coast of the United States used chaff to interfere with U.S. air defense radars.[54] Other common electromagnetic emissions, such as high-powered radio and fire control radar transmissions, can interfere with other radios and radars. Such electromagnetic interference incidents could easily appear to be deliberate hostile acts in an acute crisis.

In summary, military accidents are actions not ordered or deliberately initiated at any level in the chain of command. They are troublesome from a crisis management perspective because decision-makers may fail to realize that the accidents were unauthorized and perceive them as deliberate provocations or signals of hostile intent. In practice, however, national leaders and military commanders attempt to distinguish accidents from deliberate provocations or attacks. Military accidents include aircraft

---

[54]"Pentagon Aides Say Soviet Planes Tried to Foil U.S. Radar," New York Times, October 15, 1977, p. 2.

incidents, ship and submarine incidents, peacetime weapons incidents, wartime incidents, and other incidents. These types of military accidents do not often occur during crises, but happen often enough under other conditions that they merit being a concern in crises.

Unauthorized Deliberate Actions

Unauthorized deliberate actions are ordered or executed by tactical-level military commanders in violation of orders issued directly by national leaders, or in violation of operational guidance contained in mechanisms of indirect control. One way in which an unauthorized deliberate action can occur is for a military commander to stretch the limits on the actions he is authorized to take-- complying with a broad interpretation of the letter of his orders rather than with what he knows to be the spirit of those orders. This type of unauthorized action is especially likely when the orders given to military forces are vague or ambiguous, leaving ample room for an on-scene commander to rationalize his actions. Unauthorized deliberate actions incidents are exceedingly rare. Only a few such incidents involving U.S. forces have occurred since World War II, in every case during limited wars.

Not all unauthorized deliberate actions are harmful to crisis management efforts. An on-scene military commander with an appreciation of the political objectives being

pursued by national leaders could well decide to ignore orders that are inappropriate for the local situation and pursue a course of action that better supports crisis management efforts. Two types of unauthorized deliberate actions can be distinguished on the basis of the military commander's intentions: constructive and malicious.[55]

A constructive unauthorized action is taken in the belief that actions called for in existing orders are inappropriate under the circumstances, and that the unauthorized action would better support the national objectives in the crisis. Whether or not the outcome is constructive is a different matter, and a well-intentioned action could seriously complicate crisis management efforts. The decision made by Commander Middle East Force in January 1964 to disregard orders from Washington and have USS Manley (DD 940) evacuate American citizens from Zanzibar is an example of a constructive unauthorized action.[56] The mark of a constructive unauthorized action is an effort to inform the chain of command as soon as possible of the action taken and the reasons for taking it.

---

[55]Unauthorized deliberate actions caused by insanity on the part of a military commander will not be addressed in this study because there is extremely little empirical evidence on which to assess such incidents. The author knows of no cases, other than certain wartime battlefield atrocities, in which insanity caused an unauthorized deliberate action.

[56]See Chapter IV for a complete description of this incident.

A malicious unauthorized deliberate action is taken out of opposition to the objectives underlying specific orders, disrespect for the chain of command or the method of control being used, or frustration with particular orders felt to be unnecessarily endangering the men performing the mission. The mark of a malicious unauthorized action is an effort to conceal the action from higher authority.

The most common type of malicious unauthorized action in United States military history since World War II has been deliberate killings of non-combatant civilians by troops in the field, in violation of orders to avoid civilian casualties. The most notorious example was the March 1968 My Lai massacre. Such battlefield incidents are a product of the stress and emotion of ground combat (which were particularly intense in the guerrilla war fought in Vietnam), thus telling us little about how military commanders behave in crises.

Three examples of malicious unauthorized actions will serve to illustrate the nature of most such incidents. The first example arose from dissatisfaction with political constraints on the conduct of wartime air operations over hostile territory. During the Korean War, the United States placed strict limits on air operations near the Chinese and Soviet borders with North Korea in order to avoid incidents with China and the Soviet Union. U.S. aircraft were not permitted to enter Chinese or Soviet airspace, bombing

missions near their borders were tightly controlled from Washington, and special precautions were taken to ensure that bombs were not accidently dropped on china or the Soviet Union while attacking North Korean targets near the borders.[57] In his history of the Korean War, Joseph C. Goulden reports that a requirement for bombers to fly parallel to the border while bombing the Yalu bridges was deliberately violated by an Air Force flight commander. After concluding that the parallel approach was ineffective and unnecessarily endangered the pilots, the flight commander ordered his planes to approach perpendicular to the border. This resulted in them penetrating several miles into Chinese airspace after dropping their bombs. These violations of the bombing restrictions were successfully kept secret until long after the war.[58]

The second example of a malicious unauthorized action arose from ill-considered thrill-seeking, but essentially involved disrespect for the chain of command. On June 13, 1957, USS Hornet (CVA 16) was conducting routine flight training in the South China Sea after a port visit to Hong Kong. During this flight training, approximately eight of Hornet's pilots decided to "buzz" (make low-level passes

---

[57]Futrell, pp. 142, 208-11, 453; Field, pp. 395-6; Cagle and Manson, pp. 224-5, 243-7.

[58]Joseph C. Goulden, Korea: The Untold Story of the War (New York: McGraw-Hill, 1982), pp. 302-3.

over) the Chinese mainland near Swatow. Their motive was simple thrill-seeking (often referred to as "flat-hatting" or "skylarking"), but their actions were a clear and deliberate violation of restrictions against entering Chinese airspace. Chinese anti-aircraft guns fired on the planes, causing minor damage to a few of them, and China protested the incident. The responsible pilots initially tried to keep their action secret, but the bullet holes in some of their planes exposed their guilt. The pilots were reprimanded and the Vice Admiral responsible for the Hornet task group was relieved of command of the Seventh Fleet's carrier task force.[59]

The third example of a possibly malicious unauthorized action also arose from dissatisfaction with political constraints on the conduct of wartime air operations over hostile territory. The secret bombing of unauthorized targets in North Vietnam directed by Air Force General John D. Lavelle between November 1971 and March 1972 was out or dissatisfaction with rules of engagement strictly limiting the anti-aircraft sites that could be attacked. General Lavelle and his pilots believed--and could demonstrate with losses they suffered--that the restrictions endangered the

---

[59]"U.S. Carrier Plane Struck by Red Chinese Gunfire," New York Times, June 13, 1957, p. 1; Vice Admiral Herbert D. Riley, "The Reminiscences of Vice Admiral Herbert D. Riley, U.S. Navy (Retired)," Volume II (Annapolis, MD: U.S. Naval Institute, Oral History Program, 1972), pp. 491-4.

lives of U.S. pilots by protecting surface-to-air missiles

that were firing on aircraft flying missions over North

Vietnam. General Lavelle directed attacks on anti-aircraft

sites not authorized under the rules of engagement and

submitted false reports listing authorized targets. When

the unauthorized bombings were discovered, General Lavelle

was relieved of his command by General John D. Ryan, Air

Force Chief of Staff.[60] That General Lavelle attempted to

---

[60]U.S. House, Armed Services Committee, Investigating
Subcommittee, Unauthorized Bombing of Military Targets in
North Vietnam, Hearings, 92nd Congress, 2nd Session
(Washington, DC: U.S. Government Printing Office, 1972), pp.
7-10; Richard K. Betts, Soldiers, Statesmen, and Cold War
Crises (Cambridge, MA: Harvard University Press, 1977), pp.
49, 238-9. The evidence in this case is ambiguous. In an
interview with the author, Admiral Thomas H. Moorer,
Chairman of the Joint Chiefs of Staff at the time of the
incident, insisted that the chain of command had not--
tacitly or otherwise--authorized the bombings. Admiral
Moorer emphasized that General Lavelle was relieved of
command for lying to his superiors about the targets he was
attacking. On the other hand, the House Armed Services
Committee subcommittee that investigated the incident
expressed doubt that General Lavelle would have conducted
the secret bombings on his own, and suggested that he may
have been given tacit authorization by the military chain of
command or even top-level civilian authorities to exceed the
letter of his written orders (the rationale being that
secret bombings would be less likely to arouse political
opposition than formal expansion of the bombing). As an
aside, the dual reporting of targets used by General Lavelle
to keep the unauthorized bombings secret strongly resembles
the dual reporting used to hide the bombing of Cambodia from
March 1969 to June 1970--which had been conducted by planes
not under General Lavelle's command. The bombing of
Cambodia was ordered by President Nixon and kept secret for
reasons of international and domestic politics. Given the
Nixon Administration's well-known obsession with secrecy and
penchant for covert action, which led to its downfall, the
Lavelle case is an exceedingly poor example to cite as
evidence of military evasion of civilian control.

conceal his actions from his superiors marks this case as a malicious unauthorized action even though he was motivated by legitimate concern over the safety of his pilots.

The pattern suggested by the malicious unauthorized actions that occurred in the Korean War and Vietnam War is that political restrictions on air operations over hostile territory tend to provoke unauthorized actions. Unauthorized actions occur because the restrictions are perceived as unnecessarily endangering the lives of pilots flying missions over hostile territory, and are motivated primarily by a desire to accomplish the mission with as few pilots shot down as possible. That is an understandable desire, but actions taken in violation of political restrictions are still unauthorized. This is an example of acute tension between political and military considerations. If political leaders and military commanders are not sensitive to each other's needs, unauthorized actions are likely to occur. Such unauthorized actions do not represent military evasion of civilian control so much as a breakdown in communications between civilian and military leaders, and a resulting inability to find an acceptable compromise for managing the tensions between political and military considerations.

Incidents at Sea

The term incidents at sea is used by the U.S. Navy to designate potentially dangerous interactions between U.S.

and Soviet naval forces on and over the high seas. For this discussion, incidents at sea will be grouped into five categories: accidental and deliberate collisions, dangerous maneuvering, threatening actions and simulated attacks, and incidents between aircraft and ships.

The most dangerous incidents at sea are accidental and deliberate collisions. Accidental collisions were discussed earlier in this chapter. There were three apparently accidental collisions prior to the signing of the Incidents at Sea Agreement in 1972: On June 24, 1966, the Soviet AGI Anemometer collided with the intelligence collection ship USS Banner (AGER 1) in the Sea of Japan; on May 10 and 11, 1967, the destroyer USS Walker (DD 517), was struck by two different Soviet destroyers that had been harassing a U.S. task group in the Sea of Japan; and on May 5, 1971, the destroyer USS Hanson (DD 832) and the Soviet tug Diomede collided in the Korean Strait. The first collision resulted from aggressive intelligence collection by the Soviet AGI, the second incident occurred inadvertently during deliberate harassment, and the third was caused by carelessness on the part of the Soviet tug. In all three incidents the Soviet vessels violated the nautical rules of the road.

Dangerous maneuvering by Soviet naval vessels and naval-associated auxiliary vessels (AGIs, supply ships, tankers, etc.) was the most common type of incident at sea prior to the Incidents at Sea Agreement. Dangerous

maneuvers at close quarters, usually in violation of the nautical rules of the road, have been used by the Soviets for several purposes. The most common incidents were dangerous maneuvers by Soviet AGIs and warships conducting aggressive surveillance and intelligence collection against U.S. Navy operations. Although some of the dangerous maneuvers may have been inadvertent, most of them fit a pattern of using such maneuvers for deliberate harassment of U.S. Navy ships. In some cases such harassment appears to have had no objective other than to demonstrate Soviet contempt for the U.S. Navy--a pattern of behavior analogous to Khrushchev's attempts to intimidate the West with Soviet strategic weapons. In most cases, however, Soviet harassment was specifically intended to disrupt U.S. naval operations, such as exercises and missile tests, launch and recovery of carrier aircraft, refueling and replenishment at sea, trailing of Soviet submarines, and oceanographic survey operations. The U.S. Navy sought, within the bounds of safety, to resist Soviet intimidation at sea, resulting in frequent instances of a maritime version of the game of "chicken." To protect its aircraft carriers from dangerous maneuvers by Soviet vessels, the U.S. Navy adopted the tactic of "shouldering"--using destroyers to force Soviet vessels clear of U.S. naval formations (the Soviets used the same tactic against U.S. ships to prevent them from trailing Soviet submarines). That dangerous maneuvering by Soviet

vessels only resulted in three collisions between 1960 and 1972 is testimony to the shiphandling skills and forbearance of the U.S. Navy ship captains that had to put up with the Soviet harassment.

In 1960 Soviet naval and merchant vessels began harassing and maneuvering dangerously close to U.S. naval vessels on a regular basis. From 1960 to 1972 there were scores of serious incidents and hundreds of instances of minor harassment. The first reported serious incident occurred on April 11, 1960, when a Soviet trawler made radical maneuvers extremely close to the oceanographic survey ship USS Michelson (AGS 23) in the Norwegian Sea, fouling the U.S. ship's towed survey gear. A brief review of six of the serious incidents that occurred over the next twelve years will illustrate the nature of incidents at sea. On April 9, 1964, the Soviet merchant ship Polostsk harassed the seaplane tender USS Duxbury Bay (AVP 38) in the Red Sea, cutting across the bow of the U.S. ship at barely ten yards. On September 25-27, 1964, the Soviet ships Dzerzhiniskiy and Magmomet Glazkeyv deliberately interfered with flight operations being conducted by USS Franklin D. Roosevelt (CVA 42) in the Mediterranean. On January 10, 1965, the Soviet ship Kotelnikov interfered with underway refueling operations between the carrier USS Saratoga (CVA 60) and the oiler USS Neosho (AO 143) in the Mediterranean. On May 10-11, 1967, Soviet destroyers severely harassed the

USS _Hornet_ (CVS 16) task group while it was conducting exercises in the Sea of Japan, twice colliding with the destroyer USS _Walker_ (DD 517). On June 7-8, 1967, during the Middle East War, a Soviet Kashin-class destroyer and Mirka-class corvette harassed the USS _America_ (CVA 66) task group in the Mediterranean, almost colliding with the destroyer USS _Lawe_ (DD 763) and the _America_. On March 31, 1972, a Soviet Kotlin-class destroyer harassed USS _Sims_ (DE 1059) and USS _Pratt_ (DLG 13) while they were trailing a Soviet submarine in the Mediterranean.[61] The pattern in these incidents was deliberate Soviet harassment intended to disrupt U.S. naval operations, deliberate Soviet violations of the nautical rules of the road, and dangerous maneuvering at close quarters frequently resulting in near collisions.

The Soviets occasionally protested alleged incidents by U.S. Navy ships, but were particularly sensitive about U.S. patrol planes making low passes over Soviet ships.

---

[61] "Post-1 JAN 1960 Incidents," pp. 1-3; Sweetman, passim; U.S. Department of State, Historical Office, Bureau of Public Affairs, _American Foreign Policy: Current Documents, 1964_ (Washington, DC: U.S. Government Printing Office, 1967), pp. 669-74; "U.S. Tells Soviet It Imperils Ships," _New York Times_, April 4, 1965, p. 1; "U.S. Protests Harassment of Ships by Soviets; Rejects Soviet Charges," Department of State _Bulletin_ 52 (May 3, 1965): 655-58; "A U.S. Destroyer In Far East Bumped By Soviet Warship," _New York Times_, May 11, 1967, p. 1; "A Soviet Warship Bumps U.S. Vessel 2d Time in 2 Days," _New York Times_, May 12, 1967, p. 1; "Russians Continue to Harass 6th Fleet," _New York Times_, June 9, 1967, p. 1; "U.S. to Protect Espionage Ships," _New York Times_, June 12, 1968, p. 6. These references also describe twelve other serious incidents similar to those summarized above.

Soviet protests against U.S. naval vessels generally alleged
dangerous maneuvers close to Soviet merchant ships. The
U.S. responded to these protests by stating that the U.S.
Navy ships fully complied with the rules of the road while
maneuvering to identify the Soviet vessels.[62] Although
there undoubtedly were instances of U.S. ships and planes
being overly enthusiastic in carrying out their surveillance
missions, it is clear that the U.S. did not conduct a
sustained program of harassment against Soviet naval vessels
and merchant ships.

The Incidents at Sea Agreement was signed May 25,
1972, during the first Nixon-Brezhnev summit. The agreement
committed both sides to respect the international rules of
the road for preventing collisions at sea and provided
guidance for situations unique to naval forces (such as
formations of ships) that were not adequately covered by the
international rules. In addition to specifying behavior for
naval vessels at sea, the agreement set up a standard

---

[62]"Soviet Charges NATO Buzzing," New York Times, March
5, 1964, p. 3; "Soviet Says U.S. Buzzed Ship," New York
Times, July 9, 1964, p. 3; "Soviet Charges U.S. Buzzings,"
New York Times, April 14, 1962, p. 5; "Moscow Assails U.S.
on 'Buzzing'," New York Times, August 8, 1964, p. 6; "1,000
U.S. Ship Buzzings Charged by Soviet Paper," New York Times,
August 9, 1964, p. 5; U.S. Department of State, Historical
Office, Bureau of Public Affairs, American Foreign Policy:
Current Documents, 1964 (Washington, DC: U.S. Government
Printing Office, 1967), pp. 671-4; "U.S. Protests Harassment
of Ships by Soviets; Rejects Soviet Charges," Department of
State Bulletin 52 (May 3, 1965): 656-7; "Soviet Accuses 7th
Fleet," New York Times, March 30, 1965, p. 19.

channel for reporting violations to the other side and called for annual review of the agreement. At the first annual review, held in May 1973, a protocol to the agreement was signed that expanded its provisions.[63] The provisions of the agreement were strengthened by a 1972 revision to the international rules of the road that explicitly recognized vessels launching or recovering aircraft and vessels engaged in underway replenishment as "restricted in their ability to maneuver"--a privileged status requiring other vessels to maneuver to remain clear of them. The United States and almost all other nations had long recognized this provision (the old rules did not specify which vessels could claim this status); the Soviets had refused to do so, contributing to the large number of incidents prior to 1972.

---

[63]"Agreement Between the Government of the United States of America and the Government of the Union of Soviet Socialist Republics on the Prevention of Incidents On and Over the High Seas," Stockholm International Peace Research Institute, World Armaments and Disarmament: SIPRI Yearbook, 1973 (New York: Humanities Press, 1972), pp. 36-39. "Protocol to the Agreement Between the Government of the United States of America and the Government of the Union of Soviet Socialist Republics on the Prevention of Incidents On and Over the High Seas," signed May 25, 1973, U.S. Department of State, United States Treaties and Other International Agreements, Vol. 24, Part 1, 1973 (Washington, DC: U.S. Government Printing Office, 1974), pp. 1063-64. Also see Anthony F. Wolf, "Agreement at Sea: The United States-USSR Agreement on Incidents at Sea," Korean Journal of International Studies 9 (3/1978): 57-80; Rear Admiral Robert P. Hilton, "The U.S.-Soviet Incidents at Sea treaty," Naval Forces 6 (1/1985): 30-37; Sean M. Lynn-Jones, "The Incidents at Sea Agreement," in Alexander L. George, Philip J. Farley, and Alexander Dallin, eds., U.S.-Soviet Security Cooperation: Achievements, Failures, Lessons (New York: Oxford University Press, 1988), pp. 482-509.

Although incidents between U.S. and Soviet naval vessels declined significantly after the Incidents at Agreement was signed in 1972, they were not eliminated entirely. Secretary of the Navy John F. Lehman. Jr., stated in 1984 that the number of incidents in which there was a "potential for danger" declined from over one hundred per year in the 1960s to about forty per year in 1982-1983.[64] This 60 percent reduction in the number of serious incidents each year indicates that the Incidents at Sea Agreement has been at least a partial success. On the other hand, serious incidents continue to occur at a rate sufficient to warrant concern.

There have been three apparently accidental collisions since 1972: On August 1976, a Soviet Echo II-class nuclear-powered guided missile submarine (SSGN) collided with the frigate USS Voge (FF 1047) in the Mediterranean; on November 1983, a Soviet Krivak I-class frigate collided with the destroyer USS Fife (DD 991) in the Arabian Sea; and on March 1984, the carrier USS Kittyhawk (CV 63) collided with a Soviet Victor-class nuclear attack submarine (SSN) in the Sea of Japan.[65]

Additionally, there have been two apparently deliberate collisions: On September 4, 1973, the U.S. naval

---

[64]"U.S. Officials Weigh Protest of Sea Incident," Washington Post, April 4, 1984, p. 1.

[65]On the accidental collisions, see footnote 33.

oceanographic survey ship USNS Artemis was deliberately

rammed three times by the Soviet survey ship Nakhodka in the

Atlantic. On February 12, 1988, a Soviet Mirka-class cor-

vette deliberately collided with the destroyer USS Caron (DD

970) and, simultaneously, a Krivak-class frigate collided

with USS Yorktown (CG 48). The Soviet ships carefully

maneuvered to scrape the sides of the U.S. ships, causing

only minor damage to themselves and the U.S. ships. The two

U.S. ships were exercising the right of innocent passage

(recognized under international law) through Soviet

territorial waters off the Crimean Peninsula in the Black

Sea. There is also no doubt that the incident was

deliberate. One of the Soviet ships radioed a warning to

the U.S. ships: "Soviet ships have orders to prevent

violation of territorial waters. I am authorized to strike

your ship with one of ours."[66] There were a total of five

accidental and deliberate collisions between 1972 and 1987--

compared with only three between 1960 and 1972. The total

number of collisions and the rate at which they occur have

thus been greater since the Incidents at Sea Agreement than

they were before the agreement.[67]

---

[66]"Soviets Bump U.S. Ships in Black Sea," Washington
Post, February 13, 1988, p. A23.

[67]The rate of collisions (deliberate and accidental)
was an average of one collision every four years during the
1960-1972 period, but an average of one collision every 3.2
years during the 1972-1988 period.

Soviet naval vessels have continued the practice of maneuvering dangerously in close proximity to U.S. ships. Soviet ships attempted to interfere with the U.S. and Japanese search for the flight data and cockpit voice recorders from Korean Air Lines flight 007 in the Sea of Japan from September 2 to October 28, 1983.[68] The November 17, 1983, collision between a Soviet frigate and the USS Fife (DD 991), described above, occurred while the Soviet ship was harassing the USS Ranger (CV 61) battle group. Thus, while Soviet behavior at sea improved after the Incidents at Sea Agreement was signed, the Soviets were still willing to engage in deliberate harassment of U.S. naval operations.

Threatening actions and simulated attacks have not been eliminated by the Incidents at Sea Agreement. In fact, the most serious incident of this type occurred during the 1973 Middle East war, after the agreement was signed. On October 26, the day after the United States set DEFCON 3 worldwide, the Soviet Navy commenced intensive anti-carrier

[68]"Race to Recover 007's 'Black Box' Shapes Up at Sea and in Law Offices," Washington Post, September 17, 1983, p. A1; "On the Sea of Japan, 20 Ships Comb for the Wreckage," New York Times, September 20, 1983, p. A10; "U.S. Says Soviet Ships Harass Plane-Data Searchers," Washington Post, September 21, 1983, p. A1; "U.S. Ships Hear Flight Recorder; Soviet Is Said to Hamper Search," New York Times, September 21, 1983, p. A10; "Search Goes On for Jet's 'Black Box'," New York Times, September 23, 1983, p. A3; "U.S. Says Soviet Ship Menaced Search Boat," New York Times, October 13, 1983, p. 3.

exercises against the three U.S. carrier task groups and the U.S. amphibious task group in the eastern Mediterranean. The anti-carrier exercise consisted of simulated coordinated anti-ship missile and naval gunfire attacks against the U.S. task groups. U.S. Navy commanders were unable to distinguish the simulated attacks from real attacks until the Soviet ships pulled away without having launched missiles or fired guns. Soviet ships and sub-marines armed with anti-ship cruise missiles were constantly within range of the U.S. carriers while they were in the eastern Mediterranean, making the carriers extremely vulnerable to an actual Soviet preemptive strike.[69] The Soviet exercise, which lasted through November 3, was probably intended as a signal that the Soviet Navy was prepared to counter the Sixth Fleet in the eastern Mediterranean. Vice Admiral Murphy, Commander of the Sixth Fleet, has described the tactical situation in the Mediterranean during the Soviet

---

[69]Admiral Elmo R. Zumwalt, Jr., On Watch: A Memoir (New York: Quadrangle/New York Times Book Co., 1976), pp. 436, 447; Robert G. Weinland, "Superpower Naval Diplomacy in the October 1973 Arab-Israeli War: A Case Study," The Washington Papers, Vol. VI (Beverly Hills: Sage, 1979), p. 74; Stephen S. Roberts, "The October 1973 Arab-Israeli War," in Bradford Dismukes and James McConnell, eds., Soviet Naval Diplomacy (New York: Pergamon Press, 1979), pp. 195, 204, 206; Jon D. Glassman, Arms for the Arabs: The Soviet Union and War in the Middle East (Baltimore: Johns Hopkins University Press, 1975), p. 162; Rear Admiral James B. Morin, Commanding Officer of USS Franklin D. Roosevelt (CVA 42) during the crisis, letter to author, April 14, 1988; Rear Admiral John C. Dixon, Commanding Officer of USS John F. Kennedy (CVA 67) during the crisis, letter to author, April 18, 1988.

anti-carrier exercise: "The U.S. Sixth Fleet and the Soviet Mediterranean Fleet were, in effect, sitting in a pond in close proximity and the stage for the hitherto unlikely 'war at sea' scenario was set. This situation prevailed for several days. Both fleets were obviously in a high readiness posture for whatever might come next, although it appeared that neither fleet knew exactly what to expect."[70] Admiral Elmo R. Zumwalt, Jr., Chief of Naval Operations during the crisis, has described the period of the Soviet anti-carrier exercise in strong terms: "I doubt that major units of the U.S. Navy were ever in a tenser situation since World War II ended than the Sixth Fleet in the Mediterranean was for the week after the alert was declared."[71] If a Soviet ship or submarine had inadvertently launched an anti-ship cruise missile during the exercise, it could well have sparked an intense sea battle in the Mediterranean.[72] This tense situation at sea lasted for over a week after the Arab-Israeli cease-fire took hold and tensions in the Middle East had eased.

---

[70]Quoted in Zumwalt, On Watch, p. 447.

[71]Ibid, p. 446.

[72]The Sixth Fleet may have had intelligence indicating that the Soviets were conducting an exercise, and the Soviets may have deliberately ensured that the U.S. received that intelligence in order to avoid misunderstandings. However, an exercise can be used as cover for a preemptive attack and an inadvertent launching of an anti-ship missile could well have been misperceived as an indicator that the Soviet exercise had been operational deception.

There have been additional incidents since 1973: On September 12, 1975, a Soviet Kresta II-class cruiser trained its missile launchers and fire control radars on the carrier USS John F. Kennedy (CVA 67) in the Mediterranean. In August 1979 Soviet aircraft, including new Backfire bombers, conducted more than thirty simulated anti-ship missile attacks against the destroyers USS Caron (DD 970) and USS Farragut (DDG 6) in the Black Sea. On September 30 and October 1, 1982, Soviet Backfire bombers simulated anti-ship missile strikes against the carriers USS Enterprise (CVN 65) and USS Midway (CV 41) in the northern Pacific near the Aluetians. On February 18, 1984, a Soviet jet fighter fired its cannon into the wake of USS David R. Ray (DD 971) during a simulated attack and a Soviet helicopter passed within thirty feet of the destroyer while photographing it. This incident also occurred in the Black Sea. The danger in Soviet simulated attacks is that the actions taken by the Soviet ships and planes are valid indicators of hostile intent and grounds for firing first in anticipatory self-defense under U.S. Navy rules of engagement (See Chapter IV). U.S. navy commanders must show exemplary forbearance

---

[73]Sweetman, passim; "Soviet, in 2 Incidents, Takes U.S. Torpedo and Baits Ships," New York Times, August 11, 1979, p. 4; "New Soviet Bombers Fake Strikes Against U.S. Navy." Washington Post, November 9, 1982, p. A16; "Soviet Backfire Bombers Stalk U.S. Carrier Fleet," Washington Times, November 9, 1982, p. 1; "High Seas Diplomacy Continuing," Washington Post, June 8, 1984, p. A1.

and take what could well be grave risks in not firing while appearing to be under attack.

A much different incident occurred in 1984 in the South China Sea. On April 2 the Soviet aircraft carrier _Minsk_ fired eight flares at the frigate USS _Harold E. Holt_ (FF 1074), striking the frigate with three of the flares. The Soviet action was extremely dangerous and prohibited under the Incidents at Sea Agreement, but the U.S. frigate was not without blame in the incident. _Minsk_ had hoisted proper signals requesting the U.S. ship to stay clear, but _Holt_ had continued to make two passes by _Minsk_ at close range (about 300 yards).[74] This Soviet behavior is ironic, given the frequent and severe Soviet endangering of U.S. carriers in the past. Additionally, when contrasted with the large number of Soviet violations of the Incidents at Sea Agreement, this incident illustrates that the Soviets are capable of taking a distinctly one-sided view of the agreement: complying when it benefits them and violating the agreement when it does not. On balance, however, the Soviet Union has elected to comply with the agreement.

Incidents between aircraft and ships have also continued to occur despite the Incidents at Sea Agreement. It is routine for Soviet reconnaissance aircraft to make low

---

[74]"U.S. Ship Hit by Soviet Flares," _New York Times_, April 3, 1984, p. 3; "U.S. Officials Weigh Protest of Sea Incident," _Washington Post_, April 4, 1984, p. A1; "Moscow's Muscle Flexing," _Time_, April 16, 1984, p. 30.

passes over U.S. ships while on surveillance flights, and
routine for armed U.S. carrier-based jet fighters to
intercept and escort the Soviet planes as they approach in
order to ensure that they have peaceful intentions.
Similarly, Soviet carrier-based jet fighters (armed since
1982) routinely intercept and escort U.S. planes approaching
the Soviet carriers.[75] Such interactions occur somewhere in
the world on almost a daily basis. Normally ship
surveillance and intercept operations take place without
incident, but the Soviets occasionally violate the Incidents
at Sea Agreement. On May 15, 1979, two soviet Il-38 May ASW
patrol planes flew close by USS Midway (CV 41) at 500 feet
in altitude, forcing U.S. planes in the carrier's landing
pattern to take emergency evasive action. The U.S. filed an
Incidents at Sea Agreement protest over the incident. On
September 29 and 30, 1987, the Soviet missile range
instrumentation ship Chukotka illuminated U.S. Navy and Air
Force patrol planes with a laser, causing temporary blind-
ness in an Air Force pilot's eyes for about ten minutes.
Department of Defense spokesmen stated that Soviet ships

---

[75]"Pentagon Aides Say Soviet Planes Tried to Foil U.S.
Radar," New York Times, October 15, 1977, p. 2; "New Soviet
Bombers Fake Strikes Against U.S. Navy." Washington Post,
November 9, 1982, p. A16; "Navy: Armed Soviet Jets
Intercepted Ours," New York News, December 16, 1982, p. 3;
"Fleet Reports Soviet Air Surveillance," Washington Post,
December 20, 1984, p. A26; "Navy F14 Fighters Chase Off
Soviet Planes in Bering Sea," Washington Post, August 30,
1986, p A15.

had illuminated U.S. planes with lasers before, but this was the first instance of a pilot being affected.[76] The Incidents at Sea Agreement thus has not been totally effective in stopping dangerous Soviet actions at sea.

This review of Soviet-American incidents at sea leads to two conclusions: First, a wide range of dangerous interactions can occur when U.S. and Soviet naval forces are operating in close proximity. Incidents at sea have the potential to exacerbate superpower tensions during an acute crisis and certain incidents could provoke an outbreak of fighting if misperceived as indications of an imminent preemptive attack. Second, although the Incidents at Sea Agreement has reduced the number of the most serious incidents, it has not totally eliminated incidents at sea. The primary reason for this has been lax Soviet compliance with the agreement. There are thus ample grounds for concern that incidents between Soviet and American naval forces could seriously complicate crisis management efforts.

Conclusion

Tactical-level interactions are divided, based on the perspective of political-level decisionmakers, into two

---

[76]"Naval and Maritime Events, 1979" U.S. Naval Institute Proceedings 106 (May 1980): 58; "Light From Soviet Ship 'Disturbs' U.S. Pilot's Vision," Washington Post, October 3, 1987, p. A11; "Soviets Flashed Laser, U.S. Alleges," San Jose Mercury News, October 3, 1987, p. 1A.

major categories: deliberate military actions and inadvertent military incidents. Deliberate military actions are ordered by political-level decisionmakers. They can occur under delegated as well as direct control, and can be ordered in mechanisms of indirect control as well as directly over real-time communications links. Inadvertent military incidents are military actions that may affect the development of a crisis, but which are not specifically ordered or anticipated by national leaders. There are three categories of inadvertent military incidents: unanticipated authorized actions, military accidents, and unauthorized deliberate actions. Inadvertent military incidents are troublesome because decisionmakers may fail to realize they are unauthorized and perceive them as a deliberate provocation, signal of hostile intent, or escalation of a crisis.

This chapter used examples of inadvertent military incidents that occurred under conditions ranging from peacetime to wartime in order to define the range of incidents that could occur in a crisis. As will be seen in Chapter VII, however, military accidents occur infrequently in international crises. There are three reasons for this. First, the military chain of command normally cancels most military exercises affecting forces committed to or on standby for the crisis, greatly reducing the possibility of international incidents arising from exercise-related accidents. The primary reason why exercises are cancelled

is that the forces are needed for crisis operations, but exercises have also been cancelled to avoid potential political complications. The second reason for the rarity of crisis incidents is that the military chain of command usually advises its on-scene commanders to act with caution and to avoid provocative actions. This will be seen in all four of the crises examined in Chapter VII. The third reason for the lack of incidents in crises is best described as military prudence: on-scene commanders, motivated by self-preservation, generally avoid deliberately placing their forces in situations where they are extremely vulnerable to deliberate or inadvertent attacks. Military prudence is occasionally violated by top-level political officials ordering naval forces into dangerous waters, but on other occasions U.S. leaders have been careful to keep U.S. forces well clear of fighting in a local conflict. These three factors counteract other factors--increased tempo of operations and adversary forces in close proximity--that contribute to the occurrence of inadvertent military incidents.

The final task to be accomplished before commencing the case studies is to examine the unique features of naval operations and the perspectives that the U.S. Navy holds on crisis naval operations. This is necessary to understand the role of naval forces in crises and to assess the generality of the findings. The next chapter will examine these topics.

CHAPTER VI

NAVAL FORCE AS A POLITICAL INSTRUMENT

As was discussed in the introduction, naval forces
have characteristics that make them the type of force most
commonly favored by United States leaders for use as a
political instrument in crises. But those same
characteristics can exacerbate the problems of crisis
management. The misperception dilemma can be particularly
difficult to cope with when naval forces are used as a
political instrument. Tensions between political and
military considerations can be exacerbated due to the nature
of crisis naval operations. The crisis security dilemma is
especially acute in the naval warfare environment due to
weapons technology, tactical doctrines, and the tactical
situation created by crisis naval operations. Because naval
forces have unique operational characteristics, the first
step in researching the theory of stratified interaction
will be to explore how the theory and related concepts
presented in the previous two chapters apply to the use of
naval force as a political instrument in crises.

This chapter will first review the U.S. Navy's view of
its role as a political instrument, in order to identify the

Navy perspectives influencing employment of naval forces in crises. Second, the impact of naval forces on crisis stability will be explored, focusing on their impact on the security dilemma and the misperception dilemma. Third, the tensions between political and military considerations that arise in crises will be examined in the context of crisis naval operations.

Navy Views on Crisis Response

U.S. Navy leaders have had much to say about the peacetime role of the Navy since the end of World War II, particularly since the early 1970s. Understanding the Navy's view of its peacetime missions is important for understanding the doctrinal context within which peacetime naval missions are carried out. This doctrinal context can be described as a particular bureaucratic perspective on the use of force, but it reflects the lessons the Navy has learned over the years on the efficient and effective operation of naval forces in peacetime, and the Navy's perception of the principles and dynamics of naval warfare that would be operative in the event that fighting erupts.

The Navy, like every large organization, has an organizational philosophy or ideology which shapes and organizes the attitudes, perceptions, and thought processes of its members. Because success in combat is crucially dependent on maintaining effective command and control,

military organizations place great emphasis on formalizing their organizational philosophy.[1] This produces a wide range of formal guidance covering all aspects of military operations, from strategy to the smallest details of tactics. Doctrinal and operational guidance is incorporated into Navy standing orders, which define a broad range of operational procedures and the decision criteria used to select specific tactics or operational options in various circumstances.[2]

Given that the Navy has significant missions to perform in peacetime, the formulation of standing orders for peacetime operations is to be expected. Naval forces employed as a political instrument are guided in their actions by much more than the specific orders sending them on their mission, they are also, in most cases primarily, operating in accordance with doctrinal and operational guidance promulgated in various types of standing orders.

---

[1] See Alexander L. George, "The 'Operational Code': A Neglected Approach to the Study of Political Leaders and Decision-Making," International Studies Quarterly 13 (June 1969): 190-222; John Steinbruner, The Cybernetic Theory of Decision (Princeton, NJ: Princeton University Press, 1974), pp. 88-139; Morton Halperin, Bureaucratic Politics and Foreign Policy (Washington, DC: Brookings Institution, 1974), p. 28; Robert Jervis, Perception and Misperception in International Politics (Princeton, NJ: Princeton University Press, 1976), pp. 117-202. On organizational ideologies in military organizations, see Jack Snyder, The Ideology of the Offensive (Ithaca, NY: Cornell University Press, 1984), pp. 26-30.

[2] See Chapter IV for a detailed description of standing orders.

The guidance in those standing orders is founded on the
basic concepts used by the Navy to describe its peacetime
roles. Thus, an understanding of the Navy's views of its
missions will provide insight on how forces performing those
missions are employed.

Prior to the early 1970s the U.S. Navy did not
conceive of peacetime missions as a category separate and
distinct from wartime missions. That the Navy had peacetime
roles to perform was recognized, but, with the exception of
naval diplomacy, those roles were viewed as being derived
from wartime missions or as preparatory to execution of
wartime missions. Rear Admiral John D. Chase in 1969 listed
the functions of the Navy, in order of their historical
development, as being coastal defense, commerce raiding,
enforcing respect for U.S. interests (especially trade and
shipping), being an instrument of foreign policy, commanding
the sea, direct support of land operations, projecting force
inland from the sea, and strategic deterrence.[3] These
functions reflect the Navy's conception of its missions
during the postwar period, with the exceptions that since
the writings of Mahan coastal defense had been viewed as
best achieved by commanding the sea, and that commerce
raiding had been superseded by broader concepts of blockade

---

[3]Rear Admiral John D. Chase, "The Functions of the
Navy," U.S. Naval Institute Proceedings 95 (October 1969):
27-33. Also see Captain Daniel J. Carrison, The United
States Navy (New York: Praeger, 1968), pp. 36-55.

and control of sea lines of communication, both elements of commanding the sea.

Of the eight functions listed by Admiral Chase, three are applicable in peacetime: enforcing respect for U.S. interests, being an instrument of foreign policy, and strategic deterrence. Enforcing respect for U.S. interests entails use of force to defend against attacks on merchant shipping, diplomatic outposts, and citizens abroad. Being an instrument of foreign policy includes "showing the flag" in port visits to other countries and voyages abroad. Strategic deterrence is provided primarily by submarine launched ballistic missiles and nuclear-armed sea launched cruise missiles, although carrier aircraft can also contribute to the mission. The remaining five functions are wartime missions, their peacetime impact is that the Navy strives to maintain readiness to perform these missions in wartime. This is important, because serious conflicts can arise between performance of the peacetime functions and maintaining readiness for wartime functions. As will be seen later, this tension between peacetime missions and readiness for wartime missions is one of the most significant interactions between political and military factors affecting the use of naval forces as a political instrument.

During the tour of Admiral Elmo R. Zumwalt, Jr., as Chief of Naval Operations (CNO), serious efforts were made

to refine and clarify the Navy's conceptions of its missions. The result, as described in a 1974 article by Vice Admiral Stansfield Turner, then President of the Naval War College, was a scheme of four missions: strategic deterrence, sea control, projection of power, and naval presence.[4] Sea control and projection of power are wartime missions. The objectives of sea control are "denying the enemy the right to use some seas at some times, and asserting our own right to use some seas at some times."[5] The concept of sea control differs from the earlier concept of command of the sea in recognizing that submarines and land-based aircraft had made it virtually impossible to fully control all seas at all times. Projection of power is the use of naval force against land forces, and can take three forms, naval gunfire bombardment, strikes by carrier-based tactical aircraft or sea launched cruise missiles, and amphibious assault. Objectives include interdiction, support of troops ashore, destruction of war-making potential, and seizure of territory.[6] These two missions encompass the five wartime missions listed by Admiral Chase: sea control includes coastal defense, commerce raiding, and commanding

---

[4]Vice Admiral Stansfield Turner, "Missions of the U.S. Navy," Naval War College Review 26 (March-April 1974): 2-17.

[5]Ibid, pp. 6-10.

[6]Ibid, pp. 10-13.

the sea; while projection of power includes direct support of land operations and projecting force inland.

Strategic deterrence is both a peacetime and a wartime naval mission. Peacetime objectives of strategic deterrence are to deter all-out attack on the U.S. and its allies, to deter lesser attacks with threat of unacceptable risks, and to maintain a stable political environment in which the threat of aggression or coercion against the U.S. or its allies is minimized. The wartime objective is to deter the enemy from escalating the conflict, particularly from conventional to nuclear warfare.[7]

Naval presence is "the use of naval forces, short of war, to achieve political objectives," and has two objectives: "to deter actions inimical to the interests of the United States or its allies," and "to encourage actions that are in the interest of the United States or its allies."[8] Naval presence takes two general forms: preventive deployments, a show of force in peacetime, and reactive deployments, a show of force in response to a crisis. The primary difference between preventive and reactive deployments is that preventive deployments can rely on the implied threat of reinforcement as well as the combat capabilities on-scene to influence the situation, while

---

[7]Ibid, pp. 5-6.

[8]Ibid, pp. 14-15.

reactive deployments must rely exclusively on the combat capabilities on-scene to convey a credible threat.[9] Naval presence as defined by Admiral Turner encompasses the "instrument of foreign policy" function described by Admiral Chase. The naval presence concept was the first effort by the Navy to clarify its peacetime role as a political instrument, and continues to be important in Navy thinking today.

Commander James F. McNulty, then an instructor at the Naval War College, in 1974 provided a detailed analysis of the various political-military purposes served by naval presence. The fundamental purpose of naval presence is to "contribute to the national aim of deterring conflict." Commander McNulty identified seven specific roles for naval forces in the presence mission: (a) supporting U.S. international military commitments, such as the NATO alliance, with forward deployed forces, (b) confirming on a routine basis U.S. political commitments to other nations, by showing the flag in port visits and holding joint exercises with other navies, (c) demonstrating the capability of U.S. naval forces to act in support of national interests, (d) asserting continuing U.S. interest in important areas of the world, such as the Persian Gulf, (e) demonstrating warfighting capabilities in a tension area

---

[9]Ibid.

to deter potential opponents, and serving as an instrument
of crisis management, such as by signaling U.S. intentions,
(f) providing humanitarian aid, and (g) coercing an opponent
to comply with a preferred course of action.[10] As this list
shows, the presence mission was defined as covering the full
range of naval missions short of wartime missions.

A second study of naval presence from the same period
by Lieutenant Commander Kenneth R. McGruther, then a Naval
War College student, identified six requirements for the
naval forces employed for the presence mission: (a) the
ships should be "dear," valuable assets must be committed to
demonstrate will, (b) the warfighting capability of the
force must be impressive and proven for the political signal
to be credible, (c) the force should be multi-mission
capable for flexibility of signaling and response, (d) the
potential stay-time of the forces should be substantial from
the start to signal an intent to stay until the job is done,
(e) the fleet should be forward deployed so that forces are
readily available close to potential trouble spots, and (f)
superior command, control and communications capabilities
are essential for an effective presence role.[11] This list

[10]Commander James F. McNulty, "Naval Presence--The
Misunderstood Mission," Naval War College Review 27
(September-October 1974): 26-27.

[11]Lieutenant Commander Kenneth R. McGruther, "The Role
of Perception in Naval Diplomacy," Naval War College Review
27 (September-October 1974): 12-14.

of requirements reflects Navy thinking on the presence

mission from the early 1970s onward. Of particular interest

is that the requirements emphasize employment of highly

capable forces--high value, powerful, multi-mission, high

endurance, high connectivity assets. This approach to naval

presence raises a host of potential tensions between

performance of peacetime missions and readiness to perform

warfighting missions.

From 1972 to 1978 the Navy made only minor revisions

to its mission descriptions. In 1976 the CNO, Admiral James

L. Holloway III, reduced the number of Navy missions from

four to two (called "principle functions") by making

strategic deterrence a sub-category of power projection, and

defining naval presence to be the peacetime political impact

of capabilities for sea control and power projection in

wartime (projection of political influence through the

presence of naval power).[12] This approach did not introduce

new concepts or revise the old concepts, it merely squeezed

them into two categories--sea control and power projection--

---

[12]Admiral James L. Holloway III, "Chief of Naval
Operations Posture Statement," in U.S. Congress, Senate
Armed Services Committee, Fiscal Year 1977 Authorization for
Military Procurement, Research and Development, and Active
Duty, Selected Reserve and Civilian Personnel Strengths,
Part 2, 94th Congress, 2nd Session, 1976, pp. 1056-7. Also
see Admiral Holloway's 1977 posture statement, in U.S.
Congress, Senate Armed Services Committee, Fiscal Year 1978
Authorization for Military Procurement, Research and
Development, and Active Duty, Selected Reserve and Civilian
Personnel Strengths, Part 2, 95th Congress, 1st Session,
1977, pp. 935-938.

in order to emphasize the importance of wartime capabilities as the foundation for all naval missions.

The primary significance of the "two-mission" approach for this discussion is that it reveals an element in Navy thinking that contributes to the inherent tension between performance of peacetime missions and readiness for wartime missions. The essential, fundamental purposes of a navy are to successfully carry out its combat missions in wartime and, by extension, to maintain readiness for wartime missions during peacetime.[13] Historically, navies which have lost sight of this principle have turned out to be ineffective in wartime. Treating peacetime missions as derivative of wartime missions is thus an attempt to resolve the tension between the two categories of missions in favor of readiness to perform wartime missions.

In 1979, at the initiative of CNO Admiral Thomas B. Hayward, the U.S. Navy ceased talking in terms of missions and began emphasizing principles of naval strategy when describing its contribution to the nation's defense.[14] Gone

[13]See James A. Nathan and James K. Oliver, The Future of United States Naval Power (Bloomington: Indiana University Press, 1970), pp. 70-72, 81.

[14]Admiral Thomas B. Hayward, "The Future of U.S. Sea Power," U.S. Naval Institute Proceedings 105 (May 1979): 66-71. Admiral Hayward's shift toward strategic principles was preceded by (and undoubtedly influenced by) the Navy's Sea Plan 2000 study, completed in March 1978. This study emphasized maritime superiority, maintenance of stability with forward deployments, containment of crises with

was discussion of sea control and power projection, in its place was discussion of maritime superiority, offensive warfighting posture, and forward operations. Admiral Hayward initiated a renaissance in U.S. Navy strategic thinking. The principles he first outlined in his 1979 posture statement to Congress became the basis for the Maritime Strategy, which was formally issued in 1982.

The Maritime Strategy is the overall strategic framework guiding U.S. Navy strategic and operational planning. In the event of war with the Soviet Union, the strategy calls for offensive forward operations, seizing the initiative in the war at sea to destroy the Soviet navy and carry the war to the Soviet homeland.[15] The first phase of wartime naval operations commences as a Soviet-American crisis begins escalating toward war. Aggressive forward deployment of U.S. naval forces would commence on a global basis in order to be ready for wartime operations in strategic waters, to put the Soviet Navy on the defensive,

---

selective use of force and superior naval forces on-scene, and deterrence of global war with forces capable of defending sea-lanes, reinforcing allies, and putting pressure on the Soviets. On Sea Plan 2000 see Paul B. Ryan, First Line of Defense (Stanford, Ca.: Hoover Institution Press, 1981), pp. 128-134.

[15]Admiral James D. Watkins, "The Maritime Strategy," in The Maritime Strategy, a supplement to the U.S. Naval Institute Proceedings 112 (January 1986), pp. 4-14; Captain Linton F. Brooks, "Naval Power and National Security: The Case for the Maritime Strategy," International Security 11 (Fall 1986): 58-59, 65-69.

and to deter the Soviets from escalation.[16] Navy leaders

assert that this concept of operations is founded on battle-

proven principles of naval strategy and represents the

optimum operational scenario for successfully prosecuting a

war at sea with the Soviet Union.

The Maritime Strategy addresses the employment of

naval forces as a political instrument with greater

sophistication than any previous formulation of U.S. Navy

missions. The three non-wartime naval functions encompassed

by the strategy are deterrence, forward presence, and crisis

response.

In support of overall U.S. defense strategy, the

Maritime Strategy is primarily a deterrent strategy,

designed to deter aggression across the entire spectrum of

conflict, from terrorism to nuclear war.[17] Deterrence is

achieved through strategic nuclear deterrence patrols by

ballistic missile submarines, by maintaining a visible

forward presence demonstrating the capability and intent to

execute offensive forward operations in wartime, and by

responding to crises with credible combat capabilities to

---

[16]Watkins, "Maritime Strategy," pp. 8-10.

[17]Admiral James D. Watkins, "Posture Statement by the
Chief of Naval Operations," U.S. Congress, Senate Armed
Services Committee, Department of Defense Authorization for
Appropriation for Fiscal Year 1987, Part 3. Hearings, 99th
Congress, 2nd Session (Washington, DC: U.S. Government
Priniting Office, 1986), pp. 1087-8 (Hereafter cited as
Watkins, "Posture Statement FY-87").

deter Soviet intervention and control escalation of the conflict.[18]

As this description suggests, the Maritime Strategy emphasizes deterring the Soviets by denying them military or ions--threatening to defeat Soviet forces rather than threatening retaliation. Using Snyder's analytical scheme, emphasis in the Maritime Strategy is on deterrence by denial--altering the aggressor's estimate of the probability of gaining his objectives--though with naval strategic forces providing a threat of deterrence by punishment.[19] A strong case can be made that in many circumstances, particularly in crises located outside the immediate Soviet periphery, denial is the more effective deterrent threat. However, as will be seen later, naval forces deployed to convey denial-type deterrent threats can have an impact on perceptions much stronger than had been anticipated, and can be confronted with conflicts between immediate tasking and the demands of the warfighting operations they would have to conduct in order to deny the Soviets their objectives should deterrence fail.

The second peacetime element of the Mar_time Strategy is the routine forward deployment of U.S. naval forces in

---

[18]Watkins, "Posture Statement FY-87," pp. 1095-1100, and "Maritime Strategy," p. 8.

[19]Glenn H. Snyder, Deterrence and Defense (Princeton, NJ: Princeton University Press, 1961), pp. 14-16.

peacetime. Forward presence contributes to the credibility of the U.S. deterrent posture by demonstrating denial capabilities. Forward presence is also intended to further international stability by demonstrating support for U.S. allies and other friendly countries, thus maintaining regional balances of power. In practice, forward presence is oriented toward likely trouble spots in order to have forces readily available should fighting flare up, as well as to deter hostilities. An additional benefit is that naval forces contribute to U.S. diplomatic objectives by showing the flag in port visits.[20] Forward presence, as used in the Maritime Strategy, encompasses earlier Navy concepts of naval forces as an instrument of foreign policy (Chase) and preventive naval presence (Turner).

The third peacetime element of the Maritime Strategy is crisis response, defined as employment of naval forces to achieve specific objectives while limiting the scope of the conflict and terminating military action as soon as possible. Crisis response serves primarily to control escalation of a conflict by deterring Soviet intervention and escalatory actions by other participants. Should control of escalation not be possible, the objective of crisis response is to dominate escalation--to prevail over any threats that may arise with precise use of force, so as

---

[20]Watkins, "Posture Statement FY-87," p. 1095-1100.

to avoid increased hostilities. Naval forces have escalation control characteristics that make them well-suited for this role: mobility, readiness, flexibility, endurance, and a wide range of capabilities for precision political signaling and selective military options.[21] Emphasis in crisis response is on deterrence by denial and escalation dominance should deterrence fail. National objectives are achieved through the political impact, and, if necessary, the direct military impact, of warfighting capabilities brought to bear at the scene of a crisis.

Although the Navy's description of its peacetime roles and missions changed significantly in the early 1970s and again in the early 1980s, there are strong continuities in the perspectives underlying these changing mission formulations. Five views consistently expressed by Navy leaders are particularly important for this study. First, warfighting capabilities are the foundation for performance of peacetime missions. The ability of naval forces to deter, persuade, or impress is derived from their ability to fight. Thus, peacetime missions are non-belligerent extensions of wartime missions, or, since there is always a threat that deterrence could fail, they are pre-war precursors of wartime missions.

---

[21]Ibid, pp. 1100-1102.

Second, and closely related, peacetime missions always entail maintaining readiness to perform warfighting missions, particularly in crises. Readiness to perform warfighting missions operates on two levels: readiness of on-scene forces to engage in combat at the scene of a crisis should fighting erupt, and readiness of all operational forces, particularly forward deployed forces, to perform wartime missions should the crisis escalate to war.

The third view is that deterrence, at least below the strategic nuclear level, is achieved by denial: maintaining the capability to defeat enemy forces in battle, thus denying the enemy the ability to achieve his military objectives. Deterrence by denial applies to deterring Soviet military intervention in crises as well as to deterring adversaries in crises from aggression or escalation.

Fourth, the purposes of forward presence (presence for specific or routine political signaling) are to demonstrate denial capabilities for deterrence, and to place forces where they are available to conduct warfighting missions for denial should deterrence fail.

Fifth, the two objectives of crisis response--crisis management and escalation control--are both achieved by employing forces capable of demonstrating deterrence by denial, and, should it become necessary, capable of defeating the enemy in battle to achieve denial.

The theme underlying these five views is that combat or warfighting capabilities are the basis for conducting peacetime political missions. That Navy leaders should espouse this view is no surprise, the raison d'etre of navies being to win battles at sea. This perspective is not unique to the Navy, similar views are held by all armed forces. Beyond this, however, there is merit in military leaders focusing on readiness to perform warfighting missions, for coercive threats are by definition threats that force will be used. The key point is that military leaders and political leaders may be using the same terms with much different meanings, and viewing the same military actions as having much different purposes, in deliberations on the use of force as a political instrument. Furthermore, this is not just a problem of civil-military relations: such differences in views can arise within the military chain of command, and among political leaders.

Crisis Stability

An irony of naval crisis response is that the characteristics of naval forces that make them the preferred type of force for use as a political instrument in crises also tend to make them relatively more susceptible to crisis stability problems than other types of forces. There are three naval crisis stability problems. First, political signals sent by naval forces are especially vulnerable to

misperception, making the misperception dilemma particularly acute in naval crisis response. Second, the nature of modern naval warfare places a premium on firing first in tactical engagements, making the crisis security dilemma particularly acute in naval crisis response. Third, naval warfare may be more escalation-prone than other forms of warfare.

The first naval crisis stability problem is that the political signals sent by naval forces are especially vulnerable to misperception, making the misperception dilemma particularly acute. Virtually every study of naval diplomacy has noted the danger of the signals sent by naval forces being misperceived by the target nation or third parties. Naval officers are also aware of the problem of misperception: Admiral Turner in his article explaining the Navy's view of its presence mission pointed out that the perceptions of the country to be influenced are a factor in selecting forces for naval presence.[22]

Three primary reasons have been given for the vulnerability of signals sent by naval forces to being misperceived. First, warships, being implements of war, are inherently coercive, even when used for positive, supportive, influence-building purposes.[23] They cannot

---

[22]Turner, p. 14.

[23]Ken Booth, Navies and Foreign Policy (New York: Crane Russak, 1977), p. 27; Nathan and Oliver, p. 77.

escape their aura of menace. Thus, the signals naval forces send have coercive connotations that can serve as "noise" complicating reception of the intended signal. Second, the flexibility of naval forces, which makes them so valued by national leaders for political signaling, also makes the signals they convey inherently ambiguous. As Nathan and Oliver observe, because naval forces can be withdrawn as easily as deployed, they can signal uncertainty and lack of resolve, rather than firmness and commitment.[24] Third, naval forces send highly visible signals which can be received by a large number of countries in addition to the intended recipient. Thus, third parties can perceive signals not intentionally sent to them.[25]

The second naval crisis stability problem is that the nature of modern naval warfare places a premium on firing first in tactical engagements, making the crisis security dilemma particularly acute in naval crisis response. The nature of naval warfare is that the platforms--ships, submarines and aircraft--are fragile relative to the destructiveness of the weapons used against them. This began during the era in which guns were the main armament of ships. An individual shell hit usually could not do serious

---

[24]Nathan and Oliver, pp. 78-79. Also see Edward N. Luttwak, The Political Uses of Sea Power (Baltimore: Johns Hopkins University Press, 1974), pp. 12-13.

[25]Booth, pp. 27, 32, 42; Luttwak, p. 6.

damage, but massed gunfire could destroy a ship in short order. This led to emphasis on unilateral attrition--being able to fire on the enemy without suffering his return fire-- achieved through longer-range guns and such tactical measures as surprise and maneuver.[26]

Advent of the anti-ship cruise missile greatly exacerbated the vulnerability of platforms to weapons, allowing a single weapon to destroy a ship. Even if the missile does not sink the ship, it can knock the ship out of the battle--achieving what the Navy refers to as a "mission kill." Anti-ship missiles can be difficult to defend against, making destruction of the launch platform the most effective defense against them. U.S. Navy tactical doctrine for the defense of surface ship battle groups thus emphasizes destruction of launch platforms before they launch their missiles.[27] Soviet Navy doctrine places even

---

[26]Captain Wayne P. Hughes, Jr., Fleet Tactics: Theory and Practice (Annapolis, MD: Naval Institute Press, 1986), pp. 34-39. For an illustration of the effect of unilateral attrition, see Captain Wayne P. Hughes, Jr., "Naval Tactics and Their Influence on Strategy," Naval War College Review 39 (January-February 1986): 10-11. Also see Admiral Isaac C. Kidd, Jr., "View From the Bridge of the Sixth Fleet Flagship," U.S. Naval Institute Proceedings 98 (February 1972): 18-29; Admiral Stansfield Turner and Commander George Thibault, "Countering the Soviet Threat in the Mediterranean," U.S. Naval Institute Proceedings 103 (July 1977): 25-32.

[27]Lieutenant Commander T. Wood Parker, "Thinking Offensively," U.S. Naval Institute Proceedings 107 (April 1981): 26-31; Captain William J. Ruhe, "Antiship Missiles Launch New Tactics," U.S. Naval Institute Proceedings 108 (December 1982): 60-65; Watkins, "Maritime Strategy," p. 9.

greater emphasis on the first strike, making it a central objective of strategy as well as tactics. Soviet naval writings emphasize the importance of "the battle of the first salvo."[28] The tactical doctrines of the superpower navies interact, producing a war initiation scenario described in the U.S. Navy as the "D-day shootout."[29] The side that gets off the first salvo in the D-day shootout is likely to accrue a significant tactical advantage that could determine the outcome of the war at sea.

The technology and tactical doctrines of modern naval warfare provide conditions for crisis stability problems to arise in a crisis. Crisis stability exists when neither side has an incentive to strike the first blow, but in modern naval warfare both sides have strong tactical incentives to strike the first blow. The crisis security dilemma is that, in a crisis, many of the actions a state takes to increase its security and improve its bargaining

---

[28]Admiral of the Fleet of the Soviet Union Sergei G. Gorshkov, Red Star Rising at Sea (Annapolis, MD: Naval Institute Press, 1974), pp. 131-132; Charles D. Petersen, "About-Face in Soviet Tactics," U.S. Naval Institute Proceedings 109 (August 1983): 57-63; Lieutenant Commander Alan D. Zimm, "The First Salvo," U.S. Naval Institute Proceedings 111 (February 1985): 55-60; T.A. Fitzgerald, "Blitzkrieg at Sea," U.S. Naval Institute Proceedings 112 (January 1986): 33-38.

[29]Admiral Harry D. Train, "Decision Making and Managing Ambiguity in Politico-Military Crisis," in James G. March and Roger Weissinger-Baylon, eds., Ambiguity and Command: Organizational Perspectives on Military Decision Making (Marshfield, MA: Pitman Publishing, 1986), p. 306.

position decrease the security of the adversary. When both sides employ naval forces as a political instrument in crises, creating tactical-level interaction at the scene of the crisis, the technology and tactical doctrines of modern naval warfare almost unavoidably give rise to the crisis security dilemma. United States and Soviet naval tactical doctrines in particular emphasize the offensive and striking first in naval combat. The stratified crisis security dilemma is that, in a crisis, the security dilemma is stratified, arising from the interaction processes occurring separately at each of the three levels, and affecting the likelihood of war separately at each level. When Soviet and American naval forces are deployed to the scene of an acute crisis, the security dilemma is likely to arise at the tactical level of interaction regardless of the threat perceptions held by national leaders.

The third naval crisis stability problem is that escalation control may be more difficult in naval warfare that in other types of warfare. Several observers have expressed concern over the escalatory dangers associated with the employment of naval forces. Of particular concern to some observers is the escalatory pressure that can arise when a U.S. Navy ship is attacked. White House aide Chester Cooper, commenting on the strong Senate reaction to the 1964 Tonkin Gulf Incident, described the emotions aroused by attacks on United States ships:

>There is something very magical about an attack on an American ship on the high seas. An attack on a military base or an Army convoy doesn't stir up that kind of emotion. An attack on an American ship on the high seas is bound to set off skyrockets and the 'Star Spangled Banner' and 'Hail to the chief' and everything else.[30]

George H. Quester and Sean M. Lynn-Jones have expanded upon Cooper's remarks. Noting that "It is dreadfully dangerous to sink a major power's warship today," Quester warns that "the warships of the world have become highly prized investments, such that their loss would be likely to enrage the publics and governments that matter back home--enrage them enough to trigger off escalations that neither side might have wanted, thus setting up the deterrence and bluff mechanisms that are at the heart of 'chicken'."[31] Along the same lines, Lynn-Jones observed that "Under conditions of international tension and superpower rivalry, public opinion in a liberal democracy is likely to demand retaliation after a provocation by a major rival. Naval incidents seem to elicit particularly emotional responses in the United States." He goes on to add that "Is is, of course, relatively unlikely that a naval incident could provoke a nuclear exchange between the United States and the Soviet Union. ...An incident could, however, increase tensions and

---

[30] "The 'Phantom Battle' that Led to War," U.S. News and World Report, July 23, 1984, p. 66.

[31] George H. Quester, "Naval Armaments: The Past as Prologue," in George H. Quester, ed., Navies and Arms Control (New York: Praeger Publishers, 1980), pp. 6-7.

needlessly disrupt negotiations or other political discourse, much as the U-2 incident of 1960 forced the cancellation of the Khrushchev-Eisenhower summit."[32]

Other observers contend that there is a greater risk of nuclear war erupting at sea than ashore. This argument has been made forcefully by Desmond Ball:

> The possibility of nuclear war at sea must be regarded as at least as likely as the occurrence of nuclear war in other theaters. Indeed, there is probably a greater likelihood of accidental or unauthorized launch of sea-based nuclear weapons, and the constraints on the authorized release of nuclear weapons are possibly more relaxed than those that pertain to land-based systems. Further, there are several important factors that make it likely that any major conflict at sea would escalate to a strategic nuclear exchange relatively quickly.[33]

Incidents at sea between American and Soviet forces have been identified as a potential catalyst for the nuclear

---

[32]Sean M. Lynn-Jones, "A Quiet Success for Arms Control: Preventing Incidents at Sea," International Security 9 (Spring 1985): 164.

[33]Desmond Ball, "Nuclear War at Sea," International Security 10 (Fall 1985): 28-29. The factors Ball identifeis are the occurrence of accidents at sea, the attractiveness of ships as nuclear targets, the nuclear weapons launch autonomy of naval commanders, dual-capable wepons systems and platforms, offensive Navy anti-submarine warfare (ASW) strategy (including attacks on Soviet strategic ballistic missile submarines), incentives for Soviet preemption arising from the vulnerability of Navy ASW and command and control systems, the Navy doctrine of offensive operations in forward areas, Navy tactical nuclear weapons doctrine, Soviet doctrine for war at sea, and lack of Navy contingency planning for limiting escalation in a war at sea. Also see Barry R. Posen, "Inadvertent Nuclear War? Escalation and NATO's Northern Flank," International Security 7 (Fall 1982): 28-54; Eric J. Grove, "The Maritime Strategy and Crisis Stability," Naval Forces 8 (6/1987): 34-44.

escalation dangers described by Ball. As John Borawski notes: "The 1967 Israeli sinking [sic] of the USS Liberty, and the subsequent US uncertainty as to whether a Soviet ship had attacked the Liberty, is often cited as an example of the type of nuclear Sarajevo that could inadvertently lead to war."[34] Thus, there are at least prima facie reasons for concern that the use of naval forces as a political instrument in crises has an escalatory potential that has not been adequately addressed in studies of naval diplomacy and crisis management.

Political-Military Tensions

There are three political-military tensions: tension between political considerations and the needs of diplomatic bargaining, on the one hand, and military considerations and the needs of military operations, on the other; tension between the need for top-level control of military options in a crisis, and the need for tactical flexibility and instantaneous decision-making at the scene of the crisis; and tension between performance of crisis political missions and readiness to perform wartime combat missions. All three of these tensions are likely to arise when naval forces are used as a political instrument in a crisis.

-----

[34]John Borawski, "Risk Reduction at Sea: Naval Confidence-Building Measures," Naval Forces 3 (1/1987): 18. It must be noted that Liberty was not sunk in the attack.

The first tension is between political considerations
and the needs of diplomatic bargaining, on the one hand, and
military considerations and the needs of military opera-
tions, on the other. In his study of the political uses of
sea power, Edward N. Luttwak noted that what he termed
"suasion" (the influence effects of political signals)
operates at both a tactical level (on-scene forces) and a
political level (national-level, between states). The
implication of this, according to Luttwak, is that: "Since
men at the tactical and political levels have quite
different responsibilities, contradictions between the two
levels of suasion can be a source of acute internal
controversy, just as the conflict between tactical and
political priorities has been a chronic source of tension
between soldiers and politicians in times of war."[35] As
Luttwak suggests, the tension between political and military
objectives which can arise in a crisis is a particular
manifestation of an issue in civil-military relations
inherent in war as well as peace.

In a study of the naval presence mission of the
carrier task group led by USS Enterprise (CVN-65) during the
1971 Indo-Pakistani War, McGruther provided a good example
of this problem:

> It is presumed that the Navy will continue to play a
> primary role in reinforcing the intended perceptions

---

[35]Luttwak, p. 10.

of American intent and capability, but if the crisis managers themselves are playing for much higher stakes, it follows that they are not particularly concerned with the alternatives which are left to the opponent when a force is sent into a crisis theater. For instance, _Enterprise_ while in the Bay of Bengal was under the guns and missiles of two Soviet anticarrier warfare groups capable of destroying her by a barrage of surface-to-surface missiles before a plane could have been launched. To the national decision-makers that was not an element of the political problem and, therefore, was not the point.[36]

McGruther goes on to add, "To us in the Navy, however, it is very much the point." In this case the tension was between the self-defense needs of the naval forces on-scene, and the political objective of having those forces in a highly visible position for political signaling. The trade-off made was to pursue signaling at the cost of extreme vulnerability of the ships on-scene.

Another manifestation of the tension between political and military considerations is that military contingency plans are often inappropriate for the particular crisis at hand, requiring last-minute revision prior to being executed. In assessing the reactive mode of naval presence-- deployments made after a crisis erupts--McNulty observed that "reactive situations are usually characterized by gross

---

[36]McGruther, pp. 9-10. Rear Admiral J.R. Hill has made the same point. After explaining that political considerations may make it necessary to risk an initial casualty before the rules of engagement can be relaxed to allow prudent defensive measures, he wryly observes that "Nevertheless the command and crew of HMS _Initial Casualty_ are not likely to welcome their predicament." Hill, _Maritime Strategy for Medium Powers_ (Annapolis, MD: Naval Institute Press, 1986), p. 128.

uncertainties which require ad hoc revisions to plans on a near real-time basis. Such improvisation is a chancy business when the issue of war or peace hangs in the balance."[37] Operational plans are, of necessity, developed for specified scenarios, which may resemble the crisis at hand closely or remotely, but never anticipate it precisely. Contingency plans may be inappropriate for military as well as political reasons, but even in this case their rilitary weaknesses are likely to be the result of the manner in which the crisis developed politically.

Operational crisis management requirements that can be imposed on the use of force include limiting the size and composition of the naval force employed, placing naval forces close to or in the midst of fighting as a visible signal, limiting the actions that naval forces can take in self-defense, informing the adversary of military operations in progress against his forces (such as tracking his submarines), deliberately slowing the tempo of military operations and creating pauses in the action, and using force in gradually increasing increments. On the other hand, battle-proven principles for the successful conduct of military operations include security (keeping one's intentions secret), seizing the initiative with offensive action, surprise, concentration of superior force on the

---

[37]McNulty. p. 25.

objective, and speed in the execution of an operation and exploitation of further opportunities. The potential conflicts are obvious.[38] Because naval forces must always be prepared for the possibility of combat even while on political signaling missions, this tension between political and military considerations arises well before force is actually used.

The second tension is between the need for top-level control of military options in a crisis, and the need for tactical flexibility and instantaneous decisionmaking at the scene of the crisis. Studies of naval diplomacy and naval command and control have recognized that maintaining control of naval operations, particularly when naval forces are used as a political instrument in a crisis. As Luttwak observed: "continuous political guidance of the highest possible quality is a crucial requirement of overseas deployments: a modern oceanic fleet needs a political 'radar' as much as it needs the electronic variety."[39] The simplest, and therefore most attractive, means of ensuring such continuous political guidance is for top-level decision-makers to have direct communications with and control of on-scene forces. Supporting this view, McGruther argues for "direct and specific dialog between the crisis manager and the on-scene

---

[38]Such conflicts are discussed in Chapter III.

[39]Luttwak, p. 14.

commander," warning: "To go through a chain of command requires too much time and increases the risk of either question or answer being incorrectly understood due to oversimplification or normal relay distortion."[40] Thus, a strong case can be made for direct top-level political control of on-scene naval forces in a crisis.

Although virtually all senior military commanders recognize the need for a certain degree of direct control by top-level political authorities, there is a strong belief--particularly among naval officers--that the on-scene commander must be delegated as much authority and freedom of action as possible. Top-level decisionmakers can be overwhelmed by information overload, have insufficient time to effectively control multiple operations, and have their attention diverted by one aspect of the operations to the neglect of others. They generally do not understand the complexities of modern warfare, which can make even a small-scale operation impossible to effectively control from the White House. Communications channels often become overloaded, causing excessive delays in decisionmaking and transmission of orders to operating forces.[41]

---

[40]McGruther, pp. 14-15. Also see Rear Admiral Donald T. Poe, "Command and Control: Changeless--Yet Changing," U S. Naval Institute Proceedings 100 (October 1974): 24-25.

[41]Captain W.T.T. Pakenham, "The Command and Control of Naval Operations: Principles and Organisation," Naval Forces 7 (1/1986): 50; Beaumont, "Command Method," pp. 65-67; Creveld, p. 247-51; Poe, pp. 28-29.

Naval officers invariably believe that the on-scene commander has a superior ability to control the employment of his forces. His information about the current tactical situation is inherently superior that of his superiors. The on-scene commander requires initiative and flexibility to effectively cope with a rapidly changing tactical situation. Only the on-scene commander can effectively adapt to the inevitable "friction" in military operations--the multitude of problems that shape the execution of military plans. Centralized control of military operations can stifle initiative, weaken morale, erode authority, and cause diffusion of responsibility.[42] These are the reasons why senior naval commanders generally favor granting the on-scene commander as much freedom of action as possible.

McGruther has well described the tension arising from level of control issue:

> It is important for the task force commander to know exactly what lines the opponent cannot cross and what the appropriate responses should be if the lines are crossed. Knowing in advance what the response should

---

[42] Colonel H.A. Hadd, "Orders Firm But Flexible," U.S. Naval Institute Proceedings 88 (October 1962): 87-8; Admiral Thomas B. Hayward, "An Ex-CNO's Reflection on the Garbage Can Theory of Naval Decision Making," in James G. March and Roger Weissinger-Baylon, eds., Ambiguity and Command: Organizational Perspectives on Military Decision Making (Marshfield, MA: Pitman Publishing, 1986), p. 267; Admiral Roy L. Johnson, "The Reminiscences of Admiral Roy L. Johnson, U.S. Navy (Retired)," (Annapolis, MD: U.S. Naval Institute, Oral History Program, April 1982), pp. 182-3; Commander Linton Wells II, "Plus ca Change," U.S. Naval Institute Proceedings 111 (June 1985): 30-37; Poe, p. 28-9; Train, p. 302-3, 307.

be is a joint responsibility of the on-scene commander and the crisis manager. Left to themselves in a particular situation, the former might be likely to respond too strongly and at an earlier point in developing events; the latter is likely to prefer more restraint than a rapidly heating crisis environment may tolerate.[43]

The tactical situation can appear much different to the on-scene commander, operating under the guns of the adversary, than it does to top-level political leaders, negotiating a way out of the crisis with that same adversary. The political-military dynamics of the two levels of interaction can also be quite different, with a non-violent test of capabilities being played out on-scene as an element in a political strategy of coercive diplomacy.

The third tension is between performance of peacetime missions and readiness to perform wartime missions. Martin observed that when naval presence is exercised in an area of acute military tension, political demonstration purposes blend into preparations for warfare. That is, despite the ostensibly non-belligerent purpose of the presence mission, the naval forces must in fact have "a posture capable of accepting combat."[44] Naval forces deployed to the scene of a crisis to lend credibility to a deterrent threat are also

---

[43]McGruther, p. 10. Also see Captain Frank Andrews, "The Prevention of Surprise Attack," U.S. Naval Institute Proceedings 106 (May 1980): 134.

[44]Laurence W. Martin, The Sea in Modern Strategy (New York: Praeger, 1967), p. 143. Luttwak, p. 23, refers to this phenomenon as a "duality of mission requirements."

on-scene to take military action should deterrence fail.
They thus have two missions: to carry out their assigned
peacetime tasks, and to maintain readiness to conduct
wartime combat operations.

Although tension between performance of peacetime
missions and readiness to perform wartime missions is
inherent in crises, its impact on decision-making is a
function of the specific organizational perspectives of the
armed forces involved in the crisis. McNulty has described
the perspective commonly held by Naval Officers of presence
as opposed to the other Navy missions (deterrence, sea
control, and projection of power):

> In all instances, our naval forces are organized and
> optimized toward one or more of the other three roles,
> and their commitment to the presence mission in any
> given case must frequently conflict with their
> readiness to perform tasks in support of what is
> almost inevitably perceived as their primary mission.
> This tendency to see the presence mission as
> competitive and mutually exclusive with the remaining
> mission areas seems to pose the gravest hazard to the
> success of our Navy in support of the basic goal of
> conflict avoidance."[45]

This perspective, that wartime missions have priority over
and are the foundation for peacetime missions, was also
clearly evident in the Navy's own descriptions of its
missions, reviewed in the previous section. This view has
been consistent and strongly held for over forty years, and
remains central to Navy thinking today.

---

[45]McNulty, p. 28.

The Maritime Strategy attempts to set crisis naval operations and wartime naval operations into an overall strategic framework. Should crisis response fail and a Soviet-American crisis begin escalating toward war, the first phase of what the Maritime Strategy refers to as wartime naval operations would commence. This phase of operations is intended to be executed (and, if possible, completed) _before_ war erupts. Aggressive forward deployment of U.S. naval forces would take place on a global basis in order to deter the Soviets from launching a conventional war.[46] Again, the emphasis is on deterrence by denial, deterring the Soviets by making it clear to them that they cannot achieve their wartime aims. When this prewar deployment phase of operations commences, the tension between peacetime operations and readiness for wartime operations is resolved in favor of readiness for wartime operations.

As one would expect, given the wide range of crisis scenarios that can be envisioned, the Maritime Strategy is deliberately imprecise on when or under what circumstances the transition from peacetime crisis response to the prewar deployment phase of operations would occur. In all likelihood, though this is not stated explicitly, the two phases of the strategy would proceed simultaneously. Early transition to the prewar deployment phase of operations in a

---

[46]Watkins, "Maritime Strategy," pp. 9-11.

crisis could create serious political and crisis management problems.

Crisis management and escalation control entail much more than deterrence by denial and escalation dominance, the central strategic concepts of the Maritime Strategy. The President could well decide upon a crisis management strategy in which he is willing to accept much greater risks to U.S. naval forces than are envisioned in the Maritime Strategy. This could preclude execution of the strategy in the manner preferred by the Navy. Conversely, naval forces organized, trained, and positioned for execution of the Maritime Strategy might not be immediately responsive to unanticipated ad hoc operational requirements created by the President's crisis management strategy.

The decision to shift from crisis response to the first phase of wartime operations (prewar deployment) would undoubtedly be a momentous and difficult one for the President. He can be expected to put off making this decision for as long as possible while seeking a negotiated solution to the crisis. Equally likely is the probability that the President would order the first phase of wartime operations incrementally, to use the forward deployments as further signals of resolve and to convey increasingly strong coercive threats. This raises the question of whether or not the Navy's wartime operations plans have sufficient flexibility to allow successful conduct of wartime operations

under conditions of delayed and incremental execution of the
Maritime Strategy.

According to navy leaders, delayed or incremental
execution of wartime operations could seriously threaten the
ability of the Navy to achieve its wartime objectives.
Admiral James D. Watkins pointed this out in his 1986
description of the Maritime Strategy:

> Keys to the success of both the initial phase
> and the strategy as a whole are speed and decisiveness
> in national decisionmaking. The United States must be
> in position to deter the Soviets' "battle of the first
> salvo" or deal with that if it comes. Even though a
> substantial fraction of the fleet is forward deployed
> in peacetime, prompt decisions are needed to permit
> rapid forward deployment of additional forces in
> crisis.[47]

Admiral Watkins was arguing, in effect, for the decision to
commence the first phase of wartime operations to be made
earlier in a crisis rather than later, and decisively (all
at once) rather than incrementally. This clearly illus-
trates the nature of the tension between performance of
crisis missions and readiness to perform wartime missions.
Early and decisive execution of prewar naval deployments are
viewed by Navy leaders as crucial to the success of the
Maritime Strategy, but could well be viewed by the President
as a serious threat to crisis management. This tension is
not unique to the Maritime Strategy--it is inherent in the
use of military forces in crises.

---

[47]Ibid, p. 9.

The tension between performance of crisis political missions and readiness to perform warfighting missions also raises concerns among naval officers over political restrictions imposed in mechanisms of indirect control, particularly the rules of engagement.[48] Lieutenant Commander T. Wood Parker has expressed concern that overly restrictive rules of engagement could leave the Navy vulnerable to a pre-emptive surprise attack:

> Our specific rules of engagement, although classified and dependent on the given situation, generally require us to assume a "defensive position" and to react to a hostile act. This, of course, is not all bad, for a different type of rules might result in a miscalculation which could have catastrophic consequences. Even so, our rules of engagement put us at a disadvantage because our unit commanders and individual commanding officers are forced to think defensively prior to taking offensive action. Moreover, our present rules put us in a very unpalatable situation in that the enemy can start the war at the time and place of his choosing. Within the context of the "battle of the first salvo," so important in Soviet military thinking, our rules of engagement give the Soviet Navy a tremendous advantage. The U.S. Navy can ill-afford to absorb a massive, coordinated attack prior to being able to take offensive action.[49]

This concern arises fundamentally from the nature of modern naval warfare, in which a premium is placed on striking

---

[48] See Rear Admiral Daniel V. Gallery, The Pueblo Incident (Garden City, NY: Doubleday and Co., 1970), pp. 24-26; Lieutenant Commander Michael N. Pocalyko, "25 Years After the Blink," U.S. Naval Institute Proceedings 113 (September 1987): 43; Hill, pp. 127-8.

[49] Parker, p. 29. Also see Norman Friedman, "The Rules of Engagement Issue," in E.F. Gueritz, et al., NATO's Maritime Strategy: Issues and Developments (Washington, DC: Pergamon-Brassey's, 1987), pp. 25-32.

first. The possibility that a crisis could erupt in war
exacerbates the tension between political and military
considerations inherent in rules of engagement.

In a severe crisis, one in which Soviet-American
hostilities have risen to the point that wartime options
must start receiving consideration, the tensions between
political and military objectives becomes acute. Peacetime
political missions are prone to put naval forces in
locations other than where contingency plans for wartime
operations would have them, and can employ forces of a size
and composition other than would be optimum for wartime
operations.[50] This can have two effects. First, the naval
force carrying out the political mission may not be suitable
or available for immediate employment in wartime operations
should war break out. It could well be sunk in the first
seconds of the war, its position being well known and its
presence being an aggravation. Second, the ability of the
fleets from which the units were drawn to conduct preplanned
wartime operations can be degraded by the absence of the
units.

Efforts to minimize the impact of these effects can
entail actions which may not be compatible with the
political objectives national leaders and the diplomatic
initiatives being taken to resolve the crisis. Surging

---

[50]Train, p. 306.

ships from their homeports to replace ships pulled from
forward deployed forces to perform political mission could
be misperceived by an adversary as a signal of intent to
seek a military solution to the crisis. Using a naval force
suitable for wartime operations for a political mission, or
attempting to keep it in a location and condition of
readiness suitable for wartime operations, could cause the
same misperception. On the other hand, failure to maintain
the readiness of naval forces for wartime missions be
misperceived as signaling a lack of resolve or a willingness
to sacrifice national interests to avoid an armed clash,
thus eroding credibility, undercutting the nation's
bargaining position, and debilitating efforts to negotiate a
solution to the crisis.

Conclusion

This chapter has reviewed the U.S. Navy's view of its
role as a political instrument, examined the impact of naval
forces on crisis stability, and discussed the tensions
between political and military considerations that arise in
crisis naval operations.

Five views consistently expressed by U.S. Navy leaders
are particularly important for understanding their percep-
tions of the role of naval forces in crises. First,
warfighting capabilities are viewed as the foundation for
performance of peacetime missions. That is, peacetime

missions are viewed as non-belligerent extensions of wartime missions or as pre-war precursors of wartime missions. Second, and closely related, Navy leaders strongly believe that peacetime missions must entail maintaining readiness to perform warfighting missions, particularly in crises. Third, deterrence, at least below the strategic nuclear level, is viewed as being achieved by threat of denial: maintaining the capability to defeat enemy forces in battle, thus denying the enemy the ability to achieve his military objectives. Fourth, the purposes of forward presence are viewed as demonstrating denial capabilities for deterrence and placing forces where they are available to conduct warfighting missions should deterrence fail. Fifth, the two objectives of crisis response--crisis management and escalation control--are viewed as best achieved through employing forces capable of threatening deterrence by denial, and therefore capable of defeating the enemy in battle to achieve denial.

The theme underlying these five views is that combat or warfighting capabilities are the basis for conducting peacetime political missions. That Navy leaders should espouse this view is not surprising, as the _raison d'etre_ of navies being to win battles at sea. The key point is that political leaders and military commanders may be using the same terms with much different meanings, and viewing the same military actions as having much different purposes.

Civilian leaders may view a crisis naval deployment as serving escalation deterrence purposes, while naval leaders view it as serving escalation dominance purposes. Civilian leaders could authorize prewar naval deployments as a signal of resolve, while naval leaders execute the deployments to increase readiness for wartime operations. Such differences in perspective--civilian leaders focusing on the political considerations while military leaders focus on military considerations--can give rise to tensions between political and military considerations.

These differences in perspective are not a problem so long as the deplyments succeed in acheiving their political objectives, that is, so long as the other side refrains from escalation. Two problems could arise, however, if fighting does erupt--which could result from some sort of inadvertent incident as well as from a deliberate decision by the adversary. First, civilian leaders may not understand that by executing (either incrementally or fully) military contingency plans, they are authorizing U.S. forces to conduct combat operations under certain circumstances--such as in self-defense or anticipatory self-defense. Contingency deplyments in support of allies involved in a crisis can exacerbate thsi problem if U.S. forces are authorized to use force in support of the ally. Second, civilian leaders could unknowingly be limiting their future options to a narrow range of military operations once fighting erupts.

Viewing a particlar naval deployment as a political move, civilian leaders may not task military leaders to prepare a wide range of contingency responses to an outbreak of fighting. Military leaders always have such contingency responses, but, because their focus is on protecting their forces and ensuring victory in any engagement that might arise, the intensity and scope of combat operations they envision could well exceed what civilian leaders would have desired had they participated in the planning process. Differences in perspective can thus have serious latent implications that to not become apparent until an unanticipated incident occurs.

The characteristics of naval forces that make them the preferred type of force for use as a political instrument in crises also tend to make them relatively more susceptible to crisis stability problems than other types of forces. There are three naval crisis stability problems. First, political signals sent by naval forces are especially vulnerable to misperception, making the misperception dilemma particularly acute in naval crisis response. Second, the nature of modern naval warfare places a premium on firing first in tactical engagements, making the crisis security dilemma particularly acute in naval crisis response. Third, naval warfare may be more escalation-prone than other forms of warfare.

There are three political-military tensions: tension between political considerations and the needs of diplomatic bargaining, on the one hand, and military considerations and the needs of military operations, on the other; tension between the need for top-level control of military options in a crisis, and the need for tactical flexibility and instantaneous decision-making at the scene of the crisis; and tension between performance of crisis political missions and readiness to perform wartime combat missions. All three of these tensions are likely to arise when naval forces are employed as a political instrument in crises.

This completes the first phase of the research design, which examined three major aspects of crisis military interaction: military command and control, tactical-level military interaction, and the role of naval forces in crises. Discussion of these topics was necessary to develop specific concepts for operationalizing the theory of stratified interaction. With that task completed, the second phase of the research design can now commence. The second phase consists of four case studies of crisis naval operations, which are presented in the next chapter.